

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 07:00:00 ; Search time 3215 Seconds
(without alignments)
8560.021 Million cell updates/sec

Title: US-10-776-213-2

Perfect score: 723

Sequence: 1 cttcgcattcagcagcacac.....aaaagacataataacat 723

Scoring table: IDENTITY_NUC

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	249	34.4	434	AO874584	AO874584 V11B3 mt
2	140.8	19.5	927	CNS06GVN	AL398217 T3 end of
3	45	6.2	821	CNS0090X	AL052985 Drosophila
4	42.8	5.9	1055	CC252581	CC252581 CH261-136
5	42	5.8	1101	CNS017KX	AL108171 Drosophila
6	41.6	5.8	1101	CNS012J2	AL101624 Drosophila
7	41.4	5.7	436	BM274028	BM274028 PEST0a6
8	41.4	5.7	1101	CNS0100X	AL098379 Drosophila
9	40.8	5.6	791	AG532718	AG532718 Mus muscu
10	40.4	5.6	698	BH572702	BH572702 BGGT044TF
11	40.4	5.6	791	BI255759	BI255759 602971180
12	40.2	5.6	1000	CNS020HV	AL214492 Tetradon
13	40	5.5	366	BG628729	BG628729 cc-est1c
14	40	5.5	518	AU077905	AU077905 AU077905
15	40	5.5	845	AZ541311	AZ541311 ENT065TR
16	40	5.5	871	AZ674746	AZ674746 ENT065TR
17	39.6	5.5	179	AM477009	AM477009 g840e04.y
18	39.6	5.5	665	CA126956	CA126956 SCVPLR104
19	39.6	5.5	668	CE377922	CE377922 tigr-g8s-
20	39.6	5.5	677	CL814372	CL814372 OR_CBA003
21	39.6	5.5	716	CL814356	CL814356 OR_CBA003
22	39.6	5.5	1332	CR675673	CR675673 Tetradon
23	39.4	5.4	641	CE491438	CE491438 tigr-g8s-
24	39.2	5.4	397	BQ399446	BQ399446 NISC_mp03

C 25	39	5.4	801	7	CK461301	CK461301 931660 MA
C 26	39	5.4	1323	8	CC192318	CC192318 CH261-161
C 27	38.8	5.4	932	8	AO752307	AO752307 HS 5565_B
C 28	38.6	5.3	553	7	CN385766	CN385766 LB2TR04H2
C 29	38.6	5.3	630	8	BI9521	BI9521 T907-T7 TAM
C 30	38.6	5.3	677	9	CE498419	CE498419 tigr-g8s-
C 31	38.6	5.3	1081	9	CL510366	CL510366 SA1L_829
C 32	38.6	5.3	1101	9	CNS0037Q	AL064465 Drosophila
C 33	38.4	5.3	243	2	AM633159	AM633159 b104d09.x
C 34	38.4	5.3	594	8	AO383493	AO383493 RPI11-15
C 35	38.4	5.3	739	9	CE806390	CE806390 tigr-g8s-
C 36	38.4	5.3	1316	9	AG435289	AG435289 Mus muscu
C 37	38.2	5.3	398	2	AM133335	AM133335 bel1e08.Y
C 38	38.2	5.3	674	9	AG157962	AG157962 Pan troy1
C 39	38	5.3	600	8	BZ312600	BZ312600 i427a07.b
C 40	38	5.3	600	8	BZ327427	BZ327427 i427a07.g
C 41	38	5.3	691	9	CL951462	CL951462 OIRUD000
C 42	38	5.3	802	9	AG468867	AG468867 Mus muscu
C 43	38	5.3	914	8	AZ547087	AZ547087 ENT0X40TR
C 44	38	5.3	931	8	AZ679612	AZ679612 ENT0U83TR
C 45	38	5.3	2639	3	AF289590	AF289590 Homo sapi

ALIGNMENTS

RESULT 1
AO874584/c 434 bp DNA linear GSS 08-NOV-1999

LOCUS V11B3 mtN-3xHA/lacZ Insertion library, strain Y2278 Saccharomyces cerevisiae genomic 5', genomic survey sequence.

ACCESSION AO874584

VERSION AO874584.1 GI:6286828

KEYWORDS GSS.

SOURCE Saccharomyces cerevisiae (baker's yeast)

ORGANISM Saccharomyces cerevisiae

REFERENCE 1 (bases 1 to 434)

AUTHORS Rose-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., deBerges, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S., and Snyder, M.

Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption

Unpublished (1999)

JOURNAL Contact: Kumar A

COMMENT Unpublished (1999)

Yale University

P.O. Box 208103, New Haven, CT 06520-8103, USA

Tel: 203 432 9949

Fax: 203 432 6161

Email: anuj.kumare@yale.edu

te of mtN-3xHA/lacZ insertion.

Seq primer: GGCCTCTTCTTGGAGTAC

Class: transposon-tagged.

Location/Qualifiers

1..434

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/strain="Y2278 - S288C background, cir(0) rho(0)"

/db_xref="taxon:4932"

/lab_host="E. coli"

/clone_lib="mtN-3xHA/lacZ insertion library, strain Y2278"

/note="Vector: pHS56-Sal; A yeast genomic DNA library without 2 micron or mitochondrial DNA was prepared in pHS56-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtN-3xHA/lacZ minitransposon containing lacZ, URA3, and tetr resistance."

ORIGIN

Query Match 34.4%; Score 249; DB 8; Length 434;
 Best Local Similarity 100.0%; Pred. No. 6.2e-60;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGATTAGACGACACATCATGATGCTGCTATAAATACTACCTACGGAAA 60
 DB 311 CTTTGATTAGACGACACATCATGATGCTGCTATAAATACTACGGAAA 252
 QY 61 ACCATAAAGAGAAAGCATACCTACTTGGAAAGAAAGAGACCGCTTGTAAAGGGGAT 120
 DB 251 ACCATAAAGAGAAAGCATACCTACTTGGAAAGAAAGAGACCGCTTGTAAAGGGGAT 192
 QY 121 GGGGGCTAAGAGTCACTTCTTTTCCCTCGCGGTCGGAGCCGGACCCCTCCT 180
 DB 191 GGGGGCTAAGAGTCACTTCTTTTCCCTCGCGGTCGGAGCCGGACCCCTCCT 132
 QY 181 CTCGCCGACGATTTCTTCTTTCATATGCTTCTTATTCATCCGCTGTAAGCAAC 240
 DB 131 CTCGCCGACGATTTCTTCTTTCATATGCTTCTTATTCATCCGCTGTAAGCAAC 72
 QY 241 GCACATATGA 249
 DB 71 GCACATATGA 63

RESULT 2
 CDS06GVN 927 bp DNA linear GSS 30-NOV-2001
 LOCUS T3 end of clone AS0A00706 of library AS0A from strain CLIB 533
 DEFINITION of Saccharomyces bayanus, genomic survey sequence.
 ACCESSION AL398217 GI:12151528
 VERSION AL398217.1 GI:12151528
 KEYWORDS GSS.
 SOURCE Saccharomyces bayanus
 ORGANISM Saccharomyces bayanus
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 REFERENCE 1 (bases 1 to 927)
 AUTHORS Soulier,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
 Bolland-Fukuhara,M., Bon,E., Brotier,P., Casaregola,S.,
 de-Monigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
 Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
 Sautin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
 Wincker,P. and Weissenbach,J.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
 MEDLINE 20584711
 PUBMED 11152876
 JOURNAL 2 (bases 1 to 927)
 MEDLINE Bon,E., Neuvéglise,C., Casaregola,S., Artiguenave,F., Wincker,P.,
 Aigle,M. and Durrens,P.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 5.
 JOURNAL Saccharomyces bayanus var. uvarum
 MEDLINE FEBS Lett. 487 (1), 37-41 (2000)
 PUBMED 11152880
 JOURNAL 3 (bases 1 to 927)
 MEDLINE 20584715
 TITLE Direct Submision
 JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbicola,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES
 source
 misc_feature
 ORIGIN
 Query Match 19.5%; Score 140.8; DB 9; Length 927;
 Best Local Similarity 69.2%; Pred. No. 6.7e-29;
 Matches 321; Conservative 0; Mismatches 107; Indels 36; Gaps 8;

QY 114 GGGGAGTGGGGCTAAGAGTCACTTCTTTCCCTTCGCGGTCGGAGCCGGAGC 173
 DB 240 GCGAGAGGGGGGTAGAGAGTCACTCAC-TCTCTTCCCTTTATAGTCCGGACCGAAA 298
 QY 174 CCTCTCTCCCGCAGATTTCTTCTTCATATCTTCTTTATTCCTATCCCTTGA 233
 DB 299 CCCCCCTTCCCGCAGCGTTTCTTCGTCTT-----TTCTTTTCCCTTGG 345
 QY 224 AGCAACGACATAGCATTAATATGTCGACATCTCCAGTGTGACTGTGTATC 293
 DB 346 GACAAACGATGAGTCACTTAAGAGACGCTGCTTCTTATGCTGAT-----TGTTCGA 399
 QY 294 TCACAGTGTAAACGACACCGTGGCTCGGAAACGTTCCCTTCGTGACATTTCTAAGAACAG 353
 DB 400 AAACAGCAATAGCGACACTGGGCGCCGGAAC--GTCTCTAAGATGTTCTAAGAACAG 457
 QY 354 GGTCAAGTCTCGATAT-AGAATATTAAGCGCATTTTGTAGCGCCGCGCGC----- 406
 DB 458 GGTATATGCCCCGCAATGCTAATATATCGCGCATTTTGTAGCGCCGCGTGTGCTGC 517
 QY 407 -----GGCGCGGTTCCCAATAGGAGGCGAGTTTATCGCGAGGCTCTACTTCTCT 461
 DB 518 CGCACCGCCCTTCTCTTAATATGAGGCGAGTTATCGACATGCTCTAC-TCTTCT 576
 QY 462 ATTTGGGTAAACCCCTTCTGTTTTGCGCAGTGTGCTCGACGCTGCGCGAGAAC 521
 DB 577 ATTTGGGAACCCCTTCTGTTTTTTCGCGCAGAGGCTGCACAGGCTGCGCGAGATA 636
 QY 522 TAGTGATTAAGGATGTAAC-TTTCATGATGAGCAATTAGCAAGCG 564
 DB 637 GAGCGATTAAGGATGTGACTTTTTCATGATGAGACTTACAGAGAG 680

RESULT 3
 CDS0090X/c 821 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T873 end of BAC #
 DEFINITION BACR19021 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL052985 GI:4934433
 VERSION AL052985.1 GI:4934433
 KEYWORDS Drosophila melanogaster (fruit fly)
 SOURCE Drosophila melanogaster
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 821)
 AUTHORS Genoscope.
 TITLE Direct Submision
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

COMMENT	
Determination of this BAC-end sequence was carried out as part of collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .	
FEATURES	
SOURCE	location/Qualifiers 1..821 <code>/organism="Drosophila melanogaster"</code> <code>/mol_type="genomic DNA"</code> <code>/db_xref="taxon:7227"</code> <code>/clone="BACR19021"</code> <code>/clone_lib="RPCI-98"</code> <code>/note="end : TET3"</code>
ORIGIN	
Query Match	6.2%; Score 45; DB 9; Length 821;
Best Local Similarity	37.9%; Pred. No. 0.18;
Matches	99; Conservative 38; Mismatches 124; Indels 0; Gaps 0;
OY	462 ATTGGGAGAACCCCTTTCTGTTTTCGGCAGGTGTTCTGCGAGCTGGCGGAGAACA 521
Db	561 AYKDKACAGTCKGCTKTTTWTWKTCGCCBATATAGCCCRATTGCCCTTTTKCKCM 502
OY	522 TAGTGATAAGGATGTAACTTTCATGAGAGAAATTAGCAAGCGAAAAAATCTATGCGTA 581
Db	501 WAMTATTWMCACTATWMTTTTTTBTKSTATAADPAACAMMGASAAWWATOCACCTAAT 442
OY	582 GCTGGAGACTGTTTTCATCATATAAAAGGAGAAATTGTTGCTCATAATGTACAGTT 641
Db	441 WTGAATTAAGAGTTTTCATGCTTCTCKAAAWGCTGAAATACATWTTCKXMKKGKCGC 382
OY	642 TCCTGGAGCGCTTAACCTTTATTCGAGAGGACTATCAATCATACATATTTGCAAAA 701
Db	381 TACCCAATCTATATATWTTTTTTWTTAAAAABATATATATATTAATTAATTAATAA 322
OY	702 AAAAAGACTAATATATACA 722
Db	321 WAAAAATAAAAAATA 301
RESULT 4	
CC252581/c	1055 bp DNA linear GSS 13-MAY-2003
LOCUS	CH261-136H21.Sp6.1 CH261 Gallus gallus genomic clone CH261-136H21,
DEFINITION	genomic survey sequence.
ACCESSION	CC252581
VERSION	CC252581.1 GI:30589331
KEYWORDS	GSS.
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus. 1 (bases 1 to 1055) Kremetzki,C., Higgelbocham,J., Wyllie,K., Carter,J., McPherson,D., Warren,W., Graves,T., Mardis,E. and Wilson,R. Gallus gallus BAC End Reads Unpublished (2003) Contact: Richard K. Wilson Genome Sequencing Center Washington University School of Medicine Email: submit@wustl.edu Insert Length: 182000 Std Error: 0.00
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

FEATURES	source
Seq primer: Sp6 ATTAGGTGACACTATAG	
Class: BAC ends	
High quality sequence start: 177	
High quality sequence stop: 279.	
Location/Qualifiers	
1. 1055	
/organism="Gallus gallus"	
/mol_type="genomic DNA"	
/strain="Red Jungle Fowl"	
/db_xref="taxon:9031"	
/clone="CH261-136H21"	
/sex="Female"	
/cell_line="UCD001, inbred 256"	
/clone_1lb="CH261"	
/note="Vector: pTABAC2.1, Site_1: EcoRI; Site_2: EcoRI; CH261 Female Chicken library - for library and clone ordering information: http://www.chori.org/bacpac "	
ORIGIN	
Query Match	5.9%; Score 42.8; DB 8; Length 1055;
Best Local Similarity	55.3%; Pred. No. 0.79; Mismatches 67; Indels 0; Gaps 0;
Matches 83; Conservative 0;	
Oy	472 GCCCCTTCCTGTTTCGGCCAGTGGTGTCTGCAGAGCTGCCGCCGAGAACATAGTGTAAAG 531
Db	674 GTCCCTCTATATTTTCATATAGTGTCTGATTCCTGCAGAGCTGGAGATTAATAGTGGTGA 615
Oy	532 GGATGTAACTTTCGATGAGAGAAATTAGCAACGGAATAAACTATGCTCTAGCTGGAGTT 591
Db	614 GAGACCACTCTGAGAGGAGAAATTAATGAGAGGAGAAATATAGTGAAGAGAGATT 555
Oy	592 GTTTTCAATCATTAATAAGGAGAAATTG 621
Db	554 CTTTCCAAATAGTATAGAGAGAGATGTTG 525
RESULT 5	
CNS017KX/c	
LOCUS	
DEFINITION	CNS017KX 1101 bp DNA linear GSS 26-JUL-1999
VERSION	Drosophila melanogaster genome survey sequence Sp6 end of BAC
KEYWORDS	BACN37F10 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
SOURCE	AL108171 AL108171 GI:5628475
ORGANISM	GSS.
REFERENCE	Drosophila melanogaster (fruit fly)
AUTHORS	Eukaryotes, Metazoa: Arthropoda; Hexapoda; Insecta; Pterygota;
TITLE	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
JOURNAL	Ephydroidea; Drosophilidae; Drosophila.
COMMENT	1 (bases 1 to 1101)
	Genoscope.
	Direct Submission
	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : : BP 191 91006 EVRY cedex - FRANCE (E-mail : secrete@genoscope.cns.fr)
	- Web : www.genoscope.cns.fr
	Determination of this BAC-end and sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billand at CEPIH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES	
source	
	Location/Qualifiers
	1. 1101
	/organism="Drosophila melanogaster"
	/mol_type="genomic DNA"
	/db_xref="taxon:7227"
	/clone="BACN37F10"
	/clone_1lb="DrosBAC"
	/plasmid="pBelobAC11"
	/note="end : Sp6"


```

Db      195 AATATTGTCAGACATAATTATTCATATAATATAAATAACATAGCAAT 136
Qy      713 AATATAACT 723
Db      135 AATTATTATAT 125

RESULT 8
LOCUS   CNS0100X
DEFINITION
Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL098379.1 GI:5609990
KEYWORDS
GSS.
SOURCE  Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE
AUTHORS  Direct Submission
TITLE     Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL   BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billard at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES
source    1..1101
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACN03G04"
            /clone_lib="DrosBAC"
            /plasmid="pBelobAC11"
            /note="end : Sp6"

ORIGIN
Query Match      5.7%; Score 41.4; DB 9; Length 1101;
Best Local Similarity 20.6%; Pred. No. 2;
Matches 60; Conservative 114; Mismatches 117; Indels 0; Gaps 0;

Qy      433 AGTTTATCGGCGGAGCTTACTTCTTCCTATTTGGTAAACCCCTTCTGTTTCGACCA 492
      438 WKTDIDTWGADPDMWTDITWTWTWTWTGKMDTKTKTKMDGDTTWDKMW 497
Qy      493 GTGGTTCGACGAGCGCCGCGGAAACATAGATGAAGGATGAACCTTGCATGAGAG 552
      498 KTDKDDTKTWTWDADNAGWTDKDKTKAKKKTKTDTKTKRTGTGKKTKDGMKKTDT 557
Db      553 AATTAGCAAGCGGAAAAAATCTATGCTAGCTGCGAGTTGTTTTCATCATATAAAG 612
      558 MWGMDRABRTDGAARDRAANAKDADKADTDATKTDGWTGTTTKADTTDDTDADWMD 617
Qy      613 GAGAAATTGTTGCTCACTATGACAGTTTCGGACGCTTAACTTTATATTCAGAGA 672
      618 KDWDMKDADPDKARKKRGKGGKWKTKTKTKKTKTDKTKTCTDTRKWKDKTDKRD 677
Qy      673 CTATCAATCATACAGATTATTCGAAAAAAGACTAATTAATACAT 723
      678 GKAGDGRDKGTGTGGDSDKAKGDRAMWAKATRAAADATATAATACAT 728

RESULT 9

```

```

AGS32718
LOCUS   AGS32718 791 bp DNA linear GSS 05-JUN-2004
DEFINITION
Mus musculus molossinus DNA, clone:MSWg01-443E08.TJ, genomic survey
sequence.
ACCESSION
AGS32718
VERSION  AGS32718.1 GI:48293132
KEYWORDS
GSS.
SOURCE  Mus musculus molossinus
ORGANISM
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS  Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE     BAC end Sequences of Library MSWg01
JOURNAL   Unpublished
REFERENCE
AUTHORS  Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE     Direct Submission
JOURNAL   Submitted (17-NOV-2003) Maehira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hsp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT   Clones are derived from the mouse BAC library MSWg01. For BAC
library availability, please contact Kunya Abe (abe@rtc.riken.jp).
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyada, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TV
LIBRARY
Vector : pBAC3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
source    1..791
            /organism="Mus musculus molossinus"
            /mol_type="genomic DNA"
            /sub_species="molossinus"
            /db_xref="taxon:57486"
            /clone="MSWg01-443E08.TJ"
            /sex="male"
            /issue_type="mixture of kidney and spleen"
            /clone_lib="MSWg01 Mouse Male BAC Library"

ORIGIN
Query Match      5.6%; Score 40.8; DB 9; Length 791;
Best Local Similarity 55.6%; Pred. No. 2;
Matches 75; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy      110 GTAAGGGAGATGGGGCTAAGAATCATTCACCTTCTTTCCTTCGCGATCGGACCG 169
      290 GAAAGCGGATGATGGGGGAGAGCTTTTTCCTCCNCCGCCCCCCNCCNCC 349
Db      170 GAACCCCTCTCTCCCGGACAGATTTCTTCCTTCAATCTCTTATTTCTATCCG 229
      350 CCCCCCTTTTCCCTCCCTCTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 409
Qy      230 TTGAAGCAACCGCAC 244
      410 TTTTTCACCTTCCC 424

RESULT 10
LOCUS   BH572702 698 bp DNA linear GSS 14-DEC-2001
DEFINITION
BOGTD44TF BOGT Brassica oleracea genomic clone BOGTD44, genomic
survey sequence.
ACCESSION
BH572702
VERSION  BH572702.1 GI:17824541
KEYWORDS
GSS.

```

SOURCE	Brassica oleracea
ORGANISM	Brassica oleracea
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
AUTHORS	1 (bases 1 to 698)
TITLE	Town,C.D., Van Aken,S., Uteirpack,T., Koo,H. and Fraser,C.M.
JOURNAL	Whole genome shotgun sequencing of Brassica oleracea unpublished (2001)
COMMENT	Other_GSSs: BCGTD44TR Contact: Chris Town TIGR
FEATURES	7912 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TP Class: sheared ends.
SOURCE	Location/Qualifiers 1..698 /organism="Brassica oleracea" /mol_type="genomic DNA" /strain="TO1000DH3" /db_xref="taxon:3712" /clone="BOGTD44" /clone_1lb="BOGT" /note="Vector: pHSO1, Site 1: BstXI, 2-3 kb sheared genomic DNA inserted into pHSO1 using BstXI linkers"
ORIGIN	
Query Match	5.6%; Score 40.4; DB 8; Length 698;
Best Local Similarity	50.0%; Pred. No. 3.5;
Matches 101; Conservative	0; Mismatches 101; Indels 0; Gaps 0;
OY	518 AACATGAGTAAGAAGGATGATCACTTCGATGAGACAAATTAGCAGCGAAAAAACAATG 577
Dd	485 AAAAATTTGTTAACATCATCGTTTGCGAGGATTTAGTACTAGTAGTAATAATTGTTA 544
OY	578 GCTAGCTGGAGAGTTGTTTTCAATCATATMAAAGGAGAAAATGTGTCTCACTATGTGAC 637
Dd	545 ACTAAGTTAGGAATTAATTTTGACACATATATTGATTTTGTAGTTTGCGTTTAAATATTA 604
OY	638 AGTTCTGGAGCGTCTTAACCTTTATTTGCAGAGACTATCAATATATACAGATATGTCA 697
Dd	605 AATTAGTTTGTATGATTAATCCAAATTTGATGATTTAAAAAAAAAAAAACAATTTTTTAA 664
OY	698 AAAAAAAAAAGACTAATAATA 719
Dd	665 AAAAAATACCTACTAATATTA 686
RESULT 11	
BI255759/c	
LOCUS	BI255759 791 bp mRNA linear EST 17-JUL-2001
DEFINITION	6029771B0F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5122410 5',
ACCESSION	mRNA sequence.
VERSION	BI255759
KEYWORDS	BI255759.1 GI:14809497
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	NIH-MGC http://mgc.nci.nih.gov/ 1 (bases 1 to 791)
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: csabos-re@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Incyte Genomics, Inc.

FEATURES	SOURCE	1.	791
ORGANISM	"Homo sapiens"		
MOL TYPE	"mRNA"		
DB XREF	"taxon:9606"		
CLONE	"IMAGE:5122410"		
CLONE TYPE	"cervical carcinoma cell line"		
LAB HOST	"DH10B"		
CLONE LIB	"NIH_MGC_12"		
NOTE	"Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."		
ORIGIN			
Query Match	5.6%;	Score 40.4;	DB 4; Length 791;
Best Local Similarity	58.2%;	Pred. No. 3.6;	
Matches	71;	Conservative 0;	Mismatches 51; Indels 0; Gaps 0;
OY	588	AGTTGTTTCAATCATATAAAGGAGAAATTTGTCCTCATATGTGACAGATTCTGGG	647
DB	505	ACTCTCTCCACCTCATTAAGTCAGTGCAGATTGTCAGAGTGGGAATTCACAAAGTCAGGG	446
OY	648	ACGCTTAACTTTATTTGTCAGAGACATCAATTCATACAGATATGTGCAGAAAAA	707
DB	445	ATTCTTAATTTTGTGTTGTCAGATTATCTGAGATTAAGATTTTTAAAGAAAAA	386
OY	708	AG 709	
DB	385	TG 384	
RESULT 12			
CNS020UH		1000 bp	DNA
CNS020UH			linear
LOCUS			GSS 01-SEP-2000
DEFINITION			Tetradon nigroviridis genome survey sequence T7 end of clone 167C22 of library G from Tetradon nigroviridis, genomic survey sequence.
ACCESSION		AL214492	
VERSION		AL214492.1	GI:7873311
KEYWORDS			GSS; genome survey sequence.
SOURCE			Tetradon nigroviridis
ORGANISM			Tetradon nigroviridis
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetradon.
AUTHORS			Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizmes,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W., and Weissenbach,J.
TITLE			Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence
JOURNAL			Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE			20296633
PUBMED			10835645
REFERENCE			
AUTHORS			Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Coataz,C., Fizmes,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE			Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis
JOURNAL			Genome Res. 10 (7), 939-949 (2000)
MEDLINE			20359837
PUBMED			10899143
REFERENCE			(bases 1 to 1000)

Db 508 AAAAAAA 515

RESULT 15
 AZ541311 845 bp DNA linear GSS 14-NOV-2000
 LOCUS ENTDV65TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
 DEFINITION genomic, genomic survey sequence.
 ACCESSION AZ541311
 VERSION AZ541311 GI:11148922
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica
 ORGANISM Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.
 REFERENCE 1 (bases 1 to 845)
 AUTHORS Loftus, B., Van Aken, S. and Fraser, C.
 TITLE Determination of clone end sequences from Entamoeba histolytica
 JOURNAL HMI:IMMS sheared DNA library
 COMMENT Unpublished (2000)
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: b.loftus@tigr.org
 Clones are derived from the Entamoeba histolytica HMI:IMMS sheared
 DNA library
 Seq primer: ML3-Reverse
 Class: Shotgun
 High quality sequence start: 42
 High quality sequence stop: 612.
 Location/Qualifiers

FEATURES

source
 1..845
 /organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HMI:IMMS"
 /clone_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: PHOS1; Site 1: Bat I; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica
 using a method described by Clark and Diamond (Clark,
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
 method for isolate identification. Exp. Parasitol.
 77:450.). The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + l method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for
 whole genome shotgun sequencing projects. In Genome
 Sequencing: A Practical Approach, eds. M. Vaubin and B.
 Barrel, Oxford University Press, 1999)."

ORIGIN

Query Match 5.5%; Score 40; DB 8; length 845;
 Best Local Similarity 55.9%; Pred. No. 4.8;
 Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 588 AGTTGTTTCAATCATATAAAGGAGAAATGTTGCTACATGTGACAGTTCTGGG 647
 Db 407 AATTAAATTAATAATATCATTTTGAATATTTTATTTGTTAATGACAGAACTGAA 466
 QY 648 AGGCTTAACTTTTATGTCAGAGACTATCAATCATACAGATATGTCAAAAA 707
 Db 467 AGTTCTTAATTTAATAGTAAGAATATATATCATATGATGCTTTTCTAACA 526
 QY 708 AGACTAATATAACT 723
 Db 527 TTATGAAAAATATAT 542

Search completed: August 27, 2005, 10:01:40
 Job time : 3219 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 03:07:50 ; Search time 531 Seconds
(without alignments)
8060.213 Million cell updates/sec

Title: US-10-776-213-2
Perfect score: 723
Sequence: 1 cttcgattagcagcagcac.....aaaagactataataacat 723

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: N_Geneseq_16Dec04:*
2: geneseqn1980a:*
3: geneseqn1990a:*
4: geneseqn2000a:*
5: geneseqn2001a:*
6: geneseqn2001b:*
7: geneseqn2002a:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	723	100.0	723	4	AAD07475	Yeast pro
2	723	100.0	11427	4	AAD07497	YMR251AP
3	723	100.0	13073	4	AAD07493	YMR251AP
4	720.4	99.6	850	4	AAD07503	Yeast YMR
5	497.4	68.8	680	4	ABQ76446	S. cerevisiae
6	39	5.4	2000	8	ADAV1938	Rice gene
7	38.6	5.3	2311	6	ABBS1135	ABBS1135 cDNA enco
8	38.4	5.3	544	11	ADTF96256	ADTF96256 cDNA enco
9	37.8	5.2	7736	5	AAAS29224	AAAS29224 Genomic s
10	37.8	5.2	7736	5	ABAI6123	Human ner
11	37.8	5.2	7736	6	ABBS68364	Human DNA
12	37.8	5.2	7736	10	ADCC25486	Human CDN
13	37.4	5.2	386	5	ABV04394	Human pro
14	37	5.1	366	4	AAL35477	Human mus
15	37	5.1	366	8	ABXS8465	cDNA enco
16	37	5.1	366	12	ADJ28192	Human mus
17	37	5.1	380	6	ABX36811	Bovine ES
18	37	5.1	400	6	ABQ59188	Human col
19	36.4	5.0	201	13	ADSA41207	Human aut
20	36.4	5.0	591	13	ACNS8226	Cotton gy

21	36.4	5.0	110300	13	ADS36499	Ad36499 Human aut
22	36.2	5.0	524	10	ABT22885	Breast ca
23	36	5.0	7624	6	ABJ34113	Human imm
24	35.8	5.0	553	13	ACN62578	Cotton de
25	35.8	5.0	815	2	AAZ00806	Human sec
26	35.8	5.0	815	8	ADA39926	Human sec
27	35.8	5.0	815	8	ACC50510	Human sec
28	35.8	5.0	815	10	ADA56114	Gene enco
29	35.8	5.0	1143	8	ACA40062	ACA40062 Prokaryot
30	35.8	5.0	10517	13	ADS89438	ADS89438 Oligonuc
31	35.6	4.9	506	9	ACH17729	Human adu
32	35.6	4.9	5629	6	ABL58962	Human tum
33	35.6	4.9	5988	6	ABK09744	Human ova
34	35.6	4.9	5988	10	ADH29006	Human chr
35	35.6	4.9	5988	12	ADH2483	Human chr
36	35.6	4.9	64796	13	ACN37231	Human per
37	35.4	4.9	844	5	ABV18166	Human pro
38	35.4	4.9	33353	4	AAK70003	Human imm
39	35.4	4.9	201239	8	ACA64924	Human PL2
40	35.2	4.9	281	6	ABV96856	Human pan
41	35.2	4.9	1837	2	AAZ41383	Human nor
42	35.2	4.9	2000	8	ADA71938	Rice gene
43	35.2	4.9	2938	11	ACN88712	Breast ca
44	35.2	4.9	3197	13	ADS89390	Oligonuc
45	35.2	4.9	107330	12	ADQ97316	Mouse can

ALIGNMENTS

RESULT 1	
ID	AAD07475
AC	AAD07475; standard; DNA; 723 BP.
XX	
XX	
DT	10-AUG-2001 (first entry)
XX	
DE	Yeast promoter YMR251WA.
XX	
KW	Yeast; promoter; gene expression; fermentable carbon source; glucose;
XX	
OS	Saccharomyces cerevisiae.
XX	
PN	MO200138549-A1.
XX	
XX	31-MAY-2001.
XX	
PF	17-NOV-2000; 2000MO-SE002277.
XX	
PR	23-NOV-1999; 99SE-00004247.
XX	
PA	(ASTR) ASTRAZENCA AB.
XX	
PI	Belfield G, Oakley C;
XX	
DR	WPI, 2001-367697/38.
XX	
PT	New promoter sequences from Saccharomyces cerevisiae useful for
XX	
PT	controlling expression of homologous and heterologous nucleic acid
XX	
PS	expression in yeast cells.
XX	
XX	Claim 1, Page 67-68; 191pp; English.
XX	
CC	The invention relates to yeast promoters that are used to control the
XX	
CC	expression of homologous and heterologous nucleic acids encoding proteins
XX	
CC	and polypeptides in yeast cells. The yeast promoters are induced by a
XX	
CC	fermentable carbon source such as glucose or a non-fermentable carbon
XX	
CC	source such as ethanol or both. Therefore expression of nucleic acid
XX	
CC	molecules encoding a polypeptide under the control of the novel yeast
XX	
CC	promoters are regulated by varying the level of a fermentable carbon
XX	
CC	source or a non-fermentable carbon source or both. The yeast promoters

CC are useful for, inter alia, the high level production of proteins or
 CC polypeptides in yeast cell culture. The present DNA sequence is
 CC Saccharomyces cerevisiae YMR251WA promoter related to the invention
 XX
 XX Sequence 723 BP; 199 A; 165 C; 166 G; 193 T; 0 U; 0 Other;

Query Match 100.0%; Score 723; DB 4; Length 723;
 Best Local Similarity 100.0%; Pred. No. 9.2e-211;
 Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CTTTGATTAGACGACACATCACTAGATGCGGTGATAAATACTACGAGAAA 60
DB 1 CTTTGATTAGACGACACATCACTAGATGCGGTGATAAATACTACGAGAAA 60
QY 61 ACCATTAAGAGCAACGATACCTACTTGGAGAGAAAGAGACCGCTTGAAGGGAGAT 120
DB 61 ACCATTAAGAGCAACGATACCTACTTGGAGAGAAAGAGACCGCTTGAAGGGAGAT 120
QY 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGGACCCGGACCCCTCCT 180
DB 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGGACCCGGACCCCTCCT 180
QY 181 CTCCTCCGACGATTTCTTCTTCTATATCTCTTTATCTCCATCCGTTGAAGAAC 240
DB 181 CTCCTCCGACGATTTCTTCTTCTATATCTCTTTATCTCCATCCGTTGAAGAAC 240
QY 241 GCACTATGACTAATGCTGCTGACATCTCAATGCTGATCTGTGTATCTACAGT 300
DB 241 GCACTATGACTAATGCTGCTGACATCTCAATGCTGATCTGTGTATCTACAGT 300
QY 301 GGTAAACGCAACGCTGCTCGAAACGGTCTTCTGACAAATCTTGAACAGGGGCTACA 360
DB 301 GGTAAACGCAACGCTGCTCGAAACGGTCTTCTGACAAATCTTGAACAGGGGCTACA 360
QY 361 GGTAAACGCAACGCTGCTCGAAACGGTCTTCTGACAAATCTTGAACAGGGGCTACA 360
DB 361 GGTAAACGCAACGCTGCTCGAAACGGTCTTCTGACAAATCTTGAACAGGGGCTACA 360
QY 421 ATAGGAGCGCAGTTTATCGGCGAGCTCTAATTCTTCTTATTTGGTAAAGCCCTTTC 480
DB 421 ATAGGAGCGCAGTTTATCGGCGAGCTCTAATTCTTCTTATTTGGTAAAGCCCTTTC 480
QY 481 TGTTCGATGAGAAATTAAGAGCGGAAATTAAGAGCGGAAATTAAGAGCGGAAAT 540
DB 481 TGTTCGATGAGAAATTAAGAGCGGAAATTAAGAGCGGAAATTAAGAGCGGAAAT 540
QY 541 TTTTCGATGAGAAATTAAGAGCGGAAATTAAGAGCGGAAATTAAGAGCGGAAAT 600
DB 541 TTTTCGATGAGAAATTAAGAGCGGAAATTAAGAGCGGAAATTAAGAGCGGAAAT 600
QY 601 TCATATTAAGAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGACGCTTAACTTT 660
DB 601 TCATATTAAGAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGACGCTTAACTTT 660
QY 661 TATTGAGAGGACTATCAATCATACATATTTGTAAGAGAGAGAGAGAGAGAGAGAT 720
DB 661 TATTGAGAGGACTATCAATCATACATATTTGTAAGAGAGAGAGAGAGAGAGAT 720
QY 721 CAT 723
DB 721 CAT 723

```

RESULT 2
 AAD07497

ID AAD07497 standard; DNA; 11427 BP.

XX AAD07497;

DT 10-AUG-2001 (first entry)

DE pYMR251AP DNA plasmid.

KW Yeast; promoter; gene expression; fermentable carbon source; glucose;
 KW non-fermentable carbon source; ethanol; yeast cell culture;
 KW pYMR251AP plasmid; ds.

OS Saccharomyces cerevisiae.

OS Unidentified.

PN WO200138549-A1.

PD 31-MAY-2001.

PF 17-NOV-2000; 2000MO-SE002277.

PR 23-NOV-1999; 99SE-00004247.

PA (ASTR) ASTRAENCA AB.

PI Belfield G, Oakley C;

DR WPI; 2001-367697/38.

PT New promoter sequences from Saccharomyces cerevisiae useful for

PT controlling expression of homologous and heterologous nucleic acid

PS expression in yeast cells.

CC The invention relates to yeast promoters that are used to control the

CC expression of homologous and heterologous nucleic acids encoding proteins

CC and polypeptides in yeast cells. The yeast promoters are induced by a

CC fermentable carbon source such as glucose or a non-fermentable carbon

CC source such as ethanol or both. Therefore expression of nucleic acid

CC molecules encoding a polypeptide under the control of the novel yeast

CC promoters are regulated by varying the level of a fermentable carbon

CC source or a non-fermentable carbon source or both. The yeast promoters

CC are useful for, inter alia, the high level production of proteins or

CC polypeptides in yeast cell culture. The present sequence is pYMR251AP

CC plasmid related to the invention. This plasmid contains pPR1 and yeast

XX YMR251WA promoter

Sequence 11427 BP; 3131 A; 2589 C; 2527 G; 3180 T; 0 U; 0 Other;

Query Match 100.0%; Score 723; DB 4; Length 11427;

Best Local Similarity 100.0%; Pred. No. 3.3e-210;

Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CTTTGATTAGACGACACATCACTAGATGCGGTGATAAATACTACGAGAAA 60
DB 1 CTTTGATTAGACGACACATCACTAGATGCGGTGATAAATACTACGAGAAA 60
QY 15 CTTTGATTAGACGACACATCACTAGATGCGGTGATAAATACTACGAGAAA 74
DB 15 CTTTGATTAGACGACACATCACTAGATGCGGTGATAAATACTACGAGAAA 74
QY 61 ACCATTAAGAGCAACGATACCTACTTGGAGAGAAAGAGACCGCTTGAAGGGAGAT 120
DB 61 ACCATTAAGAGCAACGATACCTACTTGGAGAGAAAGAGACCGCTTGAAGGGAGAT 120
QY 75 ACCATTAAGAGCAACGATACCTACTTGGAGAGAAAGAGACCGCTTGAAGGGAGAT 134
DB 75 ACCATTAAGAGCAACGATACCTACTTGGAGAGAAAGAGACCGCTTGAAGGGAGAT 134
QY 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGCCGGAGCCCTCTCT 180
DB 121 GGGGGCTAAGAGTCACTTCTTTTCCCTTCCGCGGTCCGAGCCGGAGCCCTCTCT 180
QY 181 CTCCTCCGACGATTTCTTCTTCTTCAATCTTCTTATTTCTTATCCCGTTGAAGCAAC 240
DB 181 CTCCTCCGACGATTTCTTCTTCTTCAATCTTCTTATTTCTTATCCCGTTGAAGCAAC 240
QY 195 CTCCTCCGACGATTTCTTCTTCTTCAATCTTCTTATTTCTTATCCCGTTGAAGCAAC 254
DB 195 CTCCTCCGACGATTTCTTCTTCTTCAATCTTCTTATTTCTTATCCCGTTGAAGCAAC 254
QY 241 GCACTATGACTAATGCTGCTGACATCTCAATGCTGATCTGTGTATCTACAGT 300
DB 241 GCACTATGACTAATGCTGCTGACATCTCAATGCTGATCTGTGTATCTACAGT 300
QY 255 GCACTATGACTAATGCTGCTGACATCTCAATGCTGATCTGTGTATCTACAGT 314
DB 255 GCACTATGACTAATGCTGCTGACATCTCAATGCTGATCTGTGTATCTACAGT 314
QY 301 GGTAAACGCAACGCTGCTCGAAACGGTCTTCTGTCGCAATTTCTAGAACAGGGGCTACA 360
DB 301 GGTAAACGCAACGCTGCTCGAAACGGTCTTCTGTCGCAATTTCTAGAACAGGGGCTACA 360
QY 315 GGTAAACGCAACGCTGCTCGAAACGGTCTTCTGTCGCAATTTCTAGAACAGGGGCTACA 374
DB 315 GGTAAACGCAACGCTGCTCGAAACGGTCTTCTGTCGCAATTTCTAGAACAGGGGCTACA 374
QY 361 GTCTCGATTAATTAATTAAGCGCATTTTGTCTAGAGCCGCCGCCGCCGCTTTCCCA 420
DB 361 GTCTCGATTAATTAATTAAGCGCATTTTGTCTAGAGCCGCCGCCGCCGCTTTCCCA 420

```

Db 375 GTCTCGATATAGAAATTAAGGCAATTTTGTAGCGCCGCGCCGCTTTCCCA 434
 QY 421 ATAGGAGGCGGAGTTATATCGGGAGCTTACTTCTTCTATTTGGTAAAGCCCTTTC 480
 Db 435 ATAGGAGGCGGAGTTATATCGGGAGCTTACTTCTTCTATTTGGTAAAGCCCTTTC 494
 QY 481 TGTTCGCGCAGTGTGCTGCAAGCTGCGCGGAGAACATAGTATAGGATTAAC 540
 Db 495 TGTTCGCGCAGTGTGCTGCAAGCTGCGCGGAGAACATAGTATAGGATTAAC 554
 QY 541 TTTTCGATGAGAAATTAAGCAAGCGAAAAAACTATGCTAGTGGAGTTGTTTCAA 600
 Db 555 TTTTCGATGAGAAATTAAGCAAGCGAAAAAACTATGCTAGTGGAGTTGTTTCAA 614
 QY 601 TCATATTAAGGAGAAATTTGCTCCTATGTCAGTTCTGGAGCTTAACTTT 660
 Db 615 TCATATTAAGGAGAAATTTGCTCCTATGTCAGTTCTGGAGCTTAACTTT 674
 QY 661 TATTGAGAGACTATCAAAATCATACAGATATTGTCAAAAAAAAGACTAATAA 720
 Db 675 TATTGAGAGACTATCAAAATCATACAGATATTGTCAAAAAAAAGACTAATAA 734
 QY 721 CAT 723
 Db 735 CAT 737

RESULT 3

AAD07493
 ID AAD07493 standard; DNA; 13073 BP.

XX AAD07493;

DT 10-AUG-2001 (first entry)

DE PYMR251AP+Luc sequence DNA construct.

KM Yeast; promoter; gene expression; fermentable carbon source; glucose;

KM non-fermentable carbon source; ethanol; yeast cell culture;

KM PYMR251AP plasmid; luciferase gene; ds.

OS Saccharomyces cerevisiae.

OS unidentified.

OS Chimeric.

PN MO200138549-A1.

PD 31-MAY-2001.

PF 17-NOV-2000; 2000MO-SH002277.

PR 23-NOV-1999; 99SE-00004247.

PA (ASTR) ASTRAZENECA AB.

PI Belfield G, Oakley C;

DR WPI, 2001-367697/38.

XX New promoter sequences from Saccharomyces cerevisiae useful for

PT controlling expression of homologous and heterologous nucleic acid

PT expression in yeast cells.

PS Example 3; Page 88-99; 191pp; English.

CC The invention relates to yeast promoters that are used to control the

CC expression of homologous and heterologous nucleic acids encoding proteins

CC and polypeptides in yeast cells. The yeast promoters are induced by a

CC fermentable carbon source such as glucose or a non-fermentable carbon

CC source such as ethanol or both. Therefore expression of nucleic acid

CC molecules encoding a polypeptide under the control of the novel yeast

CC promoters are regulated by varying the level of a fermentable carbon

CC source or a non-fermentable carbon source or both. The yeast promoters

CC are useful for, inter alia, the high level production of proteins or
 CC polypeptides in yeast cell culture. The present sequence is PYMR251AP+Luc
 CC sequence DNA construct related to the invention. The PYMR251AP+Luc
 CC sequence contains PRB1 plasmid, luc backbone (luciferase gene) and yeast
 CC PYMR251AP promoter

SQ Sequence 13073 BP; 3584 A; 2949 C; 2934 G; 3606 T; 0 U; 0 Other;

Query Match 100.0%; Score 723; DB 4; Length 13073;

Best Local Similarity 100.0%; Pred. No. 3.5e-210; Indels 0; Gaps 0;

Matches 723; Conservative 0; Mismatches 0;

QY 1 CTTTCGATTAAGCAGCAGACATCAATAGACCTGCTCATTAATAATACATACGAAAA 60
 Db 16 CTTTCGATTAAGCAGCAGACATCAATAGACCTGCTCATTAATAATACATACGAAAA 75
 QY 61 ACCATTAAGAGCAAGCGATCTTAATTGGAAGAAAAGAGCAGCTGTAAAGGGGAT 120
 Db 76 ACCATTAAGAGCAAGCGATCTTAATTGGAAGAAAAGAGCAGCTGTAAAGGGGAT 135
 QY 121 GGGGGCTAAGAGTCACTTCTTTCCCTGCGGCTCGGACCGGGACCCCTCCT 180
 Db 136 GGGGGCTAAGAGTCACTTCTTTCCCTGCGGCTCGGACCGGGACCCCTCCT 195
 QY 181 CTCGCCGACAGATTTCTTCTTTCATATCTTCTTATCTTATCCCTGAAGCAAC 240
 Db 196 CTCGCCGACAGATTTCTTCTTTCATATCTTCTTATCTTATCCCTGAAGCAAC 255
 QY 241 GCACTATGACTAATAGTGTCTGACATCTTCAATGCTGACTTGTGTATCTCAAGT 300
 Db 256 GCACTATGACTAATAGTGTCTGACATCTTCAATGCTGACTTGTGTATCTCAAGT 315
 QY 301 GGTAAAGGAGCAGCTGCTGGAACGCTTCTTCTGGAATTTCTGAACAGGGGCTACA 360
 Db 316 GGTAAAGGAGCAGCTGCTGGAACGCTTCTTCTGGAATTTCTGAACAGGGGCTACA 375
 QY 361 GTCTCGATATAGAAATTAATAAGCGCATTTTCTAGCGCGCGCGCCGCTTCCCA 420
 Db 376 GTCTCGATATAGAAATTAATAAGCGCATTTTCTAGCGCGCGCGCCGCTTCCCA 435
 QY 421 ATAGGAGGCGGAGTTATATCGCGAGCTTACTTCTTCTTATTTGGTAAAGCCCTTTC 480
 Db 436 ATAGGAGGCGGAGTTATATCGCGAGCTTACTTCTTCTTATTTGGTAAAGCCCTTTC 495
 QY 481 TGTTCGCGCAGTGTGCTGCAAGCTGCGCGGAGAACATAGTATAGGATTAAC 540
 Db 496 TGTTCGCGCAGTGTGCTGCAAGCTGCGCGGAGAACATAGTATAGGATTAAC 555
 QY 541 TTTTCGATGAGAAATTAAGCAAGCGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 600
 Db 556 TTTTCGATGAGAAATTAAGCAAGCGAAAAAACTATGCTAGCTGGAGTTGTTTCAA 615
 QY 601 TCATATTAAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGAGCTTAACTTT 660
 Db 616 TCATATTAAGGAGAAATTTGCTCACTATGTGACAGTTTCTGGAGCTTAACTTT 675
 QY 661 TATTGAGAGACTATCAAAATCATACAGATATTGTCAAAAAAAAGACTAATAA 720
 Db 676 TATTGAGAGACTATCAAAATCATACAGATATTGTCAAAAAAAAGACTAATAA 735
 QY 721 CAT 723
 Db 736 CAT 738

RESULT 4

AAD07503
 ID AAD07503 standard; DNA; 850 BP.

XX AAD07503;

DT 10-AUG-2001 (first entry)

```

DE Yeast YMR251WA promoter region.
XX
XX Yeast; promoter; gene expression; fermentable carbon source; glucose;
KM non-fermentable carbon source; ethanol; yeast cell culture; ds.
XX
XX Saccharomyces cerevisiae.
XX
FH Key Location/Qualifiers
FT CDS 1..64
FT /tag= a
FT /product= "Yeast YMR251W open reading frame encoding
FT polypeptide"
FT complement(78..95)
FT /tag= c
FT /bound_molecly= "YMR251WA forward PCR primer"
FT /partial
FT primer_bind 786..812
FT /tag= d
FT primer_bind /bound_molecly= "YMR251WA reverse PCR primer"
FT CDS /partial
FT /tag= b
FT /product= "Yeast YMR251WA open reading frame encoding
FT polypeptide"
XX
XX WO200138549-A1.
XX
XX 31-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-SE002277.
XX
XX 23-NOV-1999; 99SE-00004247.
XX
XX (ASTR ) ASTRAZENECA AB.
XX
XX Belfield G, Oakley C;
XX
XX WPI; 2001-367697/38.
XX
XX New promoter sequences from Saccharomyces cerevisiae useful for
XX controlling expression of homologous and heterologous nucleic acid
XX expression in yeast cells.
XX
XX Example 3; Fig 14; 191pp; English.
XX
XX The invention relates to yeast promoters that are used to control the
XX expression of homologous and heterologous nucleic acids encoding proteins
XX and polypeptides in yeast cells. The yeast promoters are induced by a
XX fermentable carbon source such as glucose or a non-fermentable carbon
XX source such as ethanol or both. Therefore expression of nucleic acid
XX molecules encoding a polypeptide under the control of the novel yeast
XX promoters are regulated by varying the level of a fermentable carbon
XX source or a non-fermentable carbon source or both. The yeast promoters
XX are useful for, inter alia, the high level production of proteins or
XX polypeptides in yeast cell culture. The present DNA sequence is
XX Saccharomyces cerevisiae YMR251WA promoter region related to the
XX invention
XX
XX Sequence 850 BP; 225 A; 199 C; 193 G; 233 T; 0 U; 0 Other:
SQ
Query Match 99.6%; Score 720.4; DB 4; Length 850;
Best Local Similarity 99.9%; Pred. No. 6.2e-210;
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTTTGATTAGGACGACACATCATAGATGCGCATATAAATATACACTACGGAATA 60
Db 78 CTTTGATTAGGACGACACATCATAGATGCGCATATAAATATACACTACGGAATA 137
QY 61 ACCATTAAGAGCAAGGATACCTACTTGGAAAGAAAGAGACACGCTTGAAGGGGAT 120
Db 138 ACCATTAAGAGCAAGGATACCTACTTGGAAAGAAAGAGACACGCTTGAAGGGGAT 197
QY 121 GGGGGCTAAGAGTACTTCTTTCCCTTCCGGGCTCGAGACCCGGACCCCTCT 180

```

```

Db 198 GGGGGCTAAGAGTACTTCTTTCCCTTCCGGGCTCGAGACCCGGACCCCTCT 257
QY 181 CTTCCCGACAGATTCTCTTTCATATCTTCTTTATTCCTAATCCGTTGAACAAC 240
Db 258 CTTCCCGACAGATTCTCTTTCATATCTTCTTTATTCCTAATCCGTTGAACAAC 317
QY 241 GCACATATGATTAATATGCTGCGACATCTCGATGCGCTGATCTTGATCTTCA 300
Db 318 GCACATATGATTAATATGCTGCGACATCTCGATGCGCTGATCTTGATCTTCA 377
QY 301 GGTAAAGGACCGCTGCTCGGAAAGGTTCTTGTGACAAATTCATGAACAGGGCTCA 360
Db 378 GGTAAAGGACCGCTGCTCGGAAAGGTTCTTGTGACAAATTCATGAACAGGGCTCA 437
QY 361 GTCCTGATATTAAGATTAATTAAGCGCATTTTGTGACCGCGCGGCGCGGCTTTCC 420
Db 438 GTCCTGATATTAAGATTAATTAAGCGCATTTTGTGACCGCGCGGCGCGGCTTTCC 497
QY 421 ATAGGAGCGCGAGTTTATCGCGGAGCTTACTTCTTCTATTTGGTAAAGCCCTTTC 480
Db 498 ATAGGAGCGCGAGTTTATCGCGGAGCTTACTTCTTCTATTTGGTAAAGCCCTTTC 557
QY 481 TGTTTTCCGCAAGTGTGCTGCGAGGCTGCGCGGAGAACTAGTGAATTAAGGATTA 540
Db 558 TGTTTTCCGCAAGTGTGCTGCGAGGCTGCGCGGAGAACTAGTGAATTAAGGATTA 617
QY 541 TTTGATGAGAGATTTAGCAAGCGGAAAAAATCTAGCTGAGTGGAGTTGTTTTCA 600
Db 618 TTTGATGAGAGATTTAGCAAGCGGAAAAAATCTAGCTGAGTGGAGTTGTTTTCA 677
QY 601 TCATATTAAGGAGAAATTTGTGCTCATATGTGACAGTTTCTGGAAGCTTTA 660
Db 678 TCATATTAAGGAGAAATTTGTGCTCATATGTGACAGTTTCTGGAAGCTTTA 737
QY 661 TATTGCAAGAGCTTCAATCATACAGATATTGCTCAATTAAGGAGGATTAATTA 720
Db 738 TATTGCAAGAGCTTCAATCATACAGATATTGCTCAATTAAGGAGGATTAATTA 797
QY 721 CA 722
Db 798 AA 799

```

```

RESULT 5
AB076446
ID AB076446 standard; cDNA; 680 BP.
XX
XX AB076446;
XX
XX 21-NOV-2002 (first entry)
XX
XX S. cerevisiae BAX-associated cDNA fragment SEQ ID 317.
XX
XX Bax; Bax-resistance; cytosratic; fungicide; immunosuppressive; virucide;
XX vasotrophic; vaccine; gene therapy; proliferative disorder; cancer;
XX apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
XX neurodegeneration; cell death; ss.
XX
XX Saccharomyces cerevisiae.
XX
XX WO200264766-A2.
XX
XX 22-AUG-2002.
XX
XX 21-DEC-2001; 2001WO-EP015398.
XX
XX 22-DEC-2000; 2000EP-00870318.
XX
XX 04-JAN-2001; 2001EP-00870002.
XX
XX 09-JAN-2001; 2001EP-00870003.
XX
XX (JANC ) JANSSEN PHARM NV.
XX

```

PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX
XX MPI: 2002-667002/71.
DR P-PSDB; ABG93180.
XX
XX New isolated nucleic acid representing a synthetic BAX-gene, useful as
PT medicament for treating, preventing and/or alleviating yeast or fungal
PT infections or proliferative disorders, or for preventing apoptosis in
PT certain diseases.
XX
XX Claim 36, fig 1; 344pp; English.
XX
XX This invention describes a novel nucleic acid representing a synthetic
CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
CC resistant yeast or fungi, identifying, or obtaining and identifying Bax-
CC Candida spp. sequences that are differentially expressed in a pathway
CC eventually leading to programmed cell death or identifying inhibitors or
CC inhibitor sequences of Bax-induced cell death. The products of the
CC invention have cytostatic, fungicide, immunosuppressive, virocidic and
CC vasotropic activity and can be used in vaccines or for gene therapy. The
CC isolated nucleic acid, polypeptides, pharmaceutical compositions,
CC antisense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides, or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenic flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,
CC leukaemia, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polynucleotide associated with the Bax gene
CC described in the disclosure of the invention
XX

Sequence 680 BP; 180 A; 142 C; 166 G; 192 T; 0 U; 0 Other;

Query Match 68.8%; Score 497.4; DB 6; Length 680;
Best Local Similarity 99.8%; Pred. No. 1.3e-141;
Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 224 ATCCCGTTGAAGACCCGACCTATGACTTAATGCTGTGACATCTCCATGCTGACT 283
DB 1 ATCCCGTTGAAGACCCGACCTATGACTTAATGCTGTGACATCTCCATGCTGACT 60
QY 284 TGGTGTATCTCAAGTGTGAAGGACCGTGGCTGGAAAGCGTCTTCGTAACAATT 343
DB 61 TGTGTATCTCAAGTGTGAAGGACCGTGGCTGGAAAGCGTCTTCGTAACAATT 120
QY 344 CTAGAACAGGGGCTACAGTCTCGATATATAGATATAGGCGCATTTTGTGACGCGCC 403
DB 121 CTAGAACAGGGGCTACAGTCTCGATATATAGATATAGGCGCATTTTGTGACGCGCC 180
QY 404 GCGGCGCCCGTTTCCCAATGAGGAGCGCATTTATCGCGGAGCTCTACTTCTCTAT 463
DB 181 GCGGCGCCCGTTTCCCAATGAGGAGCGCATTTATCGCGGAGCTCTACTTCTCTAT 240
QY 464 TGGGTAAGCCCTTCTGTTTCCGCGCATGTGCTGTGACGCTGCGCGAACAATA 523
DB 241 TGGGTAAGCCCTTCTGTTTCCGCGCATGTGCTGTGACGCTGCGCGAACAATA 300
QY 524 GTGATTAAGGATTAATCTTTCATGAGATATAGCAAGCGAATAAACTATGCTAGC 583
DB 301 GTGATTAAGGATTAATCTTTCATGAGATATAGCAAGCGAATAAACTATGCTAGC 360
QY 584 TGGGAGTTGTTTTCATCATATATAAAAGGAGAAATTTGTTCTCACTATGTACAGTTTC 643
DB 361 TGGGAGTTGTTTTCATCATATATAAAAGGAGAAATTTGTTCTCACTATGTACAGTTTC 420
QY 644 TGGGAGCTTAACTTTATGAGAGACTATCAATCATATACAGATATTGTCAAAAAA 703
DB 421 TGGGAGCTTAACTTTATGAGAGACTATCAATCATATACAGATATTGTCAAAAAA 480
QY 704 AAAAGACTAATAATAACA 722
|||||

DB 481 AAAAGACTAATAATAAAA 499

RESULT 6
ADA71938
ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

KM Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; de.

OS Oryza sativa.

PN WO200300898-A1.

PD 03-JAN-2003.

PF 22-JUN-2001; 2001WO-IB001105.

PR 22-JUN-2001; 2001WO-IB001105.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;

XX MPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 5.4%; Score 39; DB 8; Length 2000;
Best Local Similarity 8.1%; Pred. No. 0.66;
Matches 45; Conservative 258; Mismatches 248; Indels 2; Gaps 1;

QY 75 AGGATACCTACTTGGAGAGAAAGAGACCGCTTGAAGGGAGATGGGCGCTAAGAACT 134
DB 184 AYYKSGSMKRMWMSCGSGGGRSAVSRYGTSTKRYTKYKMYTYSASRCRAYMTTS 243
QY 135 CATTCACCTTCTTTCCCTCGCGGTCGGAGCCCGGACCCCTCTCCCGACAGATT 194
DB 244 YSMACSSYTWCSKRSMMKMKMRKRSRGWYSMSYKMMCTAYKKSYSRMCYMR 303
QY 195 TCTTCCTTCATATCTCTCTTATTCCTATCCGTTGAAGCAACGACATGACTATAA 254
DB 304 GGGWGTATYWGVRGWSRMAAMMYKXWYVYRGKMKRGWAGRMMRSKCRWSKACYVR 363
QY 255 TGTGTGACACATCTCCATGCGTGTGACTTGTGTATCTCACAGTGGTAAGGACCGCT 314
DB 364 WMMWRTRRRRWAKKSSRTSRKXKWKMRKXKXKMGYSRMSRCKRARMMKCRSGNA 423
:::|||||

OS Homo sapiens.
XX US2003087818-A1.
XX
XX 08-MAY-2003.
XX
XX 01-FEB-2002; 2002US-00066543.
XX
XX 02-FEB-2001; 2001US-0267400P.
XX 07-FEB-2001; 2001US-0267382P.
XX 11-MAY-2001; 2001US-0290322P.
XX 12-JUL-2001; 2001US-0305265P.
XX 16-AUG-2001; 2001US-0313077P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secret H,
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX WPI; 2003-040540/03.
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1775; 87pp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridizes to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage,
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
XX Sequence 544 BP; 181 A; 70 C; 73 G; 216 T; 0 U; 4 Other;
SQ
Query Match 5.3%; Score 38.4; DB 11; Length 544;
Best Local Similarity 57.5%; Pred. No. 0.55; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 601 TCATATATAAGGAGAAATGTTGCTCCTATGTGACAGTTTGGAGCTTAACTTT 660
DB 126 TCATTGTAATGTGTGAATTCATATACAGGCTATGTAATAATTTTAACTAATGCTATTTT 67
QY 661 TATTGACAGGACTATCAATTCATACAGATTTGTCAAAAAAGAGCTAATAATA 720
DB 66 GAAAAAATAAATTAATAAATAATCAATCAAAAAAATAAATAAATAAATAAATAAATAA 7
RESULT 9
ID AAS29224/c
XX AAS29224 standard; DNA: 7736 BP.
XX
XX AAS29224;
XX

DT 21-NOV-2001 (first entry)
XX
XX Genomic sequence #67 encoding novel human DNA-binding protein.
DE
XX Human; DNA-binding protein; histone; chromo domain protein;
XX Chromatin organisation modifier; Y-box binding protein; DNA organisation;
XX gene transcription; malignant disease; autoimmune disorder;
XX rheumatic disease; genetic abnormality; infectious disease;
XX neurological disorder; gene therapy; immunomodulatory; anti-HIV;
XX anti rheumatic; anti microbial; cytoskeletal; ds.
XX
XX Homo sapiens.
XX
XX WO200155162-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001305.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0224513P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225277P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226686P.
XX 22-AUG-2000; 2000US-0227009P.
XX 23-AUG-2000; 2000US-0227182P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229345P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232398P.


```
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239835P.
PR 13-OCT-2000; 2000US-0239837P.
PR 20-OCT-2000; 2000US-0240360P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251898P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM,
PI WPI; 2001-465557/50.
XX
XX Nucleic acid molecules encoding human secreted chromosomal binding
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers.
XX
XX Disclosure; SEQ ID NO 333; 561bp; English.
XX
XX The present invention relates to the isolation of novel DNA-binding
CC proteins (AAU18154-AAU18281), and cDNA and genomic sequences encoding for
CC these proteins. DNA-binding proteins such as histones, chromo (chromatin
CC organisation modifier) domain proteins, and Y-box binding proteins may
CC contribute to diseases resulting from aberrant DNA organisation and/or
CC gene transcription. The sequences of the invention are useful in
CC screening assays to identify antagonists and/or agonists that may enhance
CC or block activities mediated by DNA-binding proteins. Blockers of DNA-
CC binding proteins may be useful in treating disorders such as malignant
CC diseases (e.g. cancer), autoimmune disorders (e.g. diabetes mellitus),
CC rheumatic diseases (e.g. rheumatoid arthritis), genetic abnormalities
CC (e.g. cystic fibrosis), infectious diseases (e.g. HIV) and neurological
CC disorders (e.g. Alzheimer's disease). The polynucleotide sequences of the
CC invention may also be used in gene therapy. AA829158-AA829239 represent
CC genomic sequences encoding for novel DNA-binding proteins. Note: the
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 7736 BP; 1672 A; 1981 C; 2164 G; 1919 T; 0 U; 0 Other;

Query Match 5.2%; Score 37.8; DB 5; Length 7736;
Best Local Similarity 54.7%; Pred. No. 2.9;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 110 GTAAGGGGATGGGGCTAAGATCATCTTCTTTCCCTCGGGTCCGACCCG 169
DB 5184 GTTCAGGGGATGGATGATTAAGACACACAGATGTTCCCCACAGCCGCAATGTG 5125
QY 170 GGACCCCTCTCTCCCGCAGATTTCTTCTTCATATCTTCTTTATTCCTATCCCG 229
DB 5124 GAAGTCTTCACCTCTCTCCCGCAGTTCCTTTCCCTCATAGGCTCTAAGCTCGTCCCG 5065
QY 230 TTGAAGCAACCGCACTA 246
DB 5064 TGGTAGCAGCTGTACCA 5048

RESULT 10
ABA16123/c
ID ABA16123 standard; DNA; 7736 BP.
XX
XX ABA16123;
AC
XX 23-JAN-2002 (first entry)
DT
XX
XX Human nervous system related polynucleotide SEQ ID NO 8454.
DE
XX
XX Human; nootropic; neuroprotective; cyostatic; dermatological; vitruide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisticking; antianaemic; antirheumatic; cancer;
```


KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; da.

OS Homo sapiens.

XX WO200159063-A2.

XX 16-AUG-2001.

PF 17-JAN-2001; 2001WO-US001334.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180665P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234979P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235634P.
PR 27-SEP-2000; 2000US-0235636P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249307P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
DR Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
PS Disclosure; SEQ ID NO 8454; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/publ/published_pct_sequences
XX
SQ Sequence 7736 BP; 1672 A; 1981 C; 2164 G; 1919 T; 0 U; 0 Other;
XX
XX
XX Query Match 5.2%; Score 37.8; DB 5; Length 7736;
XX Best Local Similarity 54.7%; Pred. No. 2.9;
XX Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
XX
OY 110 GTAAAGGGGATGGGGGCTAAGAAGTCATTCCTTCTTCCCTTCGGGTCGGAGCCG 169
DB 5184 GTTCAGGGATGATGATTAAGACACACAGTGTTCCTCCCGACGCCCAAGTGG 5125
OY 170 GGACCCCTCTCTCCCGCAGCATTTCTTCAATATCTTCTTATTCCTAATCCG 229
DB 5124 GAAGTACTCAGTCTCTCCGAGTCTGCTTCCCTAATGCGCTGTGACTGCTCCG 5065
OY 230 TTGAAGCAACCGCACTA 246
DB 5064 TGGTAGCAGCTGTACCA 5048
XX
XX
XX RESULT 11
XX ABS68364/c
XX ID ABS68364 standard; DNA; 7736 BP.
XX
XX ABS68364;
XX
XX 18-NOV-2002 (first entry)
XX
XX
XX Human DNA-binding protein genomic DNA sequence #67.
XX
XX
XX Human; DNA-binding protein; B cell immunodeficiency; autoimmune disorder;
KW severe combined immunodeficiency; rheumatoid arthritis; Crohn's disease;
KW diabetes mellitus; allergy; asthma; inflammatory condition; thrombosis;
KW graft-versus-host disease; blood-related disorder; atherosclerosis;
KW hyperproliferative disorder; cancer; renal disorder; arrhythmia;
KW acute glomerulonephritis; cardiovascular disorder; respiratory disorder;
KW Goodpasture's syndrome; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; endocrine disorder; Addison's disease;
KW reproductive system disorder; endometriosis; infectious disease;
KW viral infection; bacterial infection; fungal infection; vaccine;
XX

KW gastrointestinal disorder; multiple sclerosis; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX US2002102638-A1.
XX
PD 01-AUG-2002.
XX
XX
XX 17-JAN-2001; 2001US-00764846.
XX
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-690611/74.
XX
XX Novel DNA-binding protein useful for diagnosis, prognosis, prevention and
PT treatment of immune, hyperproliferative, respiratory, cardiovascular,
PT reproductive, endocrine, gastrointestinal and neurological disorders.
XX

PS Claim 1; SEQ ID NO 333; 225bp; English.
XX
CC The present invention relates to a new DNA-binding protein. The invention
CC is useful in treating, preventing, diagnosing and/or prognosing
CC immunodeficiencies (e.g. B cell immunodeficiencies, severe combined
CC immunodeficiencies), autoimmune disorders (rheumatoid arthritis, multiple
CC sclerosis, diabetes mellitus), allergic reactions and conditions (e.g.
CC asthma), inflammatory conditions, graft-versus-host disease, blood-
CC related disorders (thrombosis, atherosclerosis), hyperproliferative
CC disorders (e.g. cancer), renal disorders (e.g. acute glomerulonephritis),
CC cardiovascular disorders (e.g. arrhythmia), respiratory disorders
CC (Goodpasture's syndrome), neurological disorders (e.g. Alzheimer's
CC disease, Parkinson's disease), endocrine disorders (e.g. Addison's
CC disease), reproductive system disorders (e.g. endometriosis), infectious
CC diseases (e.g. viral, bacterial or fungal infections) and
CC gastrointestinal disorders (e.g. Crohn's disease). The invention is also
CC useful to stimulate neuronal growth and treat, prevent, and/or diagnose
CC neuronal damage which occurs in certain neuronal disorders or neuro-
CC degenerative conditions. The present nucleic acid sequence represents a
CC human DNA-binding protein genomic DNA sequence of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at <http://seqdata.uspto.gov/sequence>
XX
SQ Sequence 7736 BP; 1672 A; 1981 C; 2164 G; 1919 T; 0 U; 0 Other;
Query Match 5.2%; Score 37.8; DB 6; Length 7736;
Best Local Similarity 54.7%; Pred. No. 2.9;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 110 GTAAGGGGATGCGGGCTAAGATGATTCATCTTCTTCCCTGCGGCGGACCCG 169
DB 5184 GTTCAGGGGATGATGTAAGACACACACAGTGTTCCTCCCGACGCCAGATGTC 5125
QY 170 GAACCCCTCTCTCCCGCACGATTCCTTCCTTCAATCTCTTATCTTATCCCG 229
DB 5124 GAAGTACTCCACTCTCTCCCGAGTCGCTTCCCTCATGAGCCTCGACTCGTCCCC 5065
QY 230 TTGAAGCAACCGACCTA 246
DB 5064 TGGTAGCAGCTGACCA 5048
RESULT 12
ADC25486/C
ID ADC25486 standard; cDNA; 7736 BP.
XX
AC ADC25486;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human cDNA from extracellular matrix gene 78 #3.
XX
XX Extracellular matrix protein; cytosolic; antibacterial; virucide;
XX neuroprotective; gynaecological; gastrointestinal; cardiac;
XX cardiovascular; Gen; nephrotoxic; anti-inflammatory; muscular; Gen;
XX respiratory; Gen; immunosuppressive; cerebroprotective; vasotropic;
XX neotropic; anti-allergic; cancer; bacterial infection; viral infection;
XX neural disorder; immune system disorder; blood disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; cardiovascular disorder; renal disorder;
XX inflammatory disorder; proliferative disorder; human; gene therapy; ss;
XX gene.
XX
XX Homo sapiens.
OS
XX
PN US2003049650-A1.
XX
PD 13-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00091483.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR

PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 15-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225577P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0233968P.
PR 12-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-023398P.
PR 14-SEP-2000; 2000US-023399P.
PR 14-SEP-2000; 2000US-0233999P.
PR 14-SEP-2000; 2000US-023400P.
PR 14-SEP-2000; 2000US-023401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR

PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764846.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Ruben SM, Barash SC,
XX
XX MPI; 2003-605749/57.
DR P-PSDB; ADC25241.
XX
XX
PT New DNA-binding proteins and gene encoding them, useful for diagnosing,
PT treating and/or preventing e.g. neurological, inflammatory, infectious,

PT cardiovascular, autoimmune, respiratory, neoplastic or digestive
PT diseases.
XX
XX
PS Disclosure; SEQ ID NO 333; 226pp; English.
XX
XX
CC The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human extracellular matrix protein, representing one of 161
CC novel genes. Also included are recombinant vectors, host cells
CC (expressing the protein), the extracellular matrix proteins (including
CC their fragments, epitopes and homologues), an isolated antibody that
CC binds specifically to the protein, diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or absence of a mutation in the nucleic acid and diagnosing a
CC condition based on the presence or absence of the mutation), diagnosing a
CC pathological condition or susceptibility to a pathological condition
CC (comprising determining the presence or amount of expression of the
CC protein in a biological sample and diagnosing a condition based on the
CC presence or amount of expression of the protein), preventing, treating or
CC ameliorating a medical condition by administering the nucleic acid or
CC protein to a mammalian subject, identifying a binding partner to the
CC protein, the gene corresponding to the cDNA sequence, and identifying an
CC activity in a biological assay (comprising expressing the nucleic acid in
CC a cell, isolating the supernatant, detecting an activity in a biological
CC assay and identifying the protein in the supernatant having the
CC activity). The nucleic acids and proteins display the following
CC activities Cytostatic, antibacterial, Virucide, Neuroprotective,
Query Match 5.2%; Score 37.8; DB 10; Length 7736;
Best Local Similarity 54.7%; Pred. No. 2.9;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 110 GTAAGGGGATGGGGCTAAGATGATTCCTTTCTTCGCGGTCGGAGCCG 169
Db 5184 GTTCAGGGATGGATGTAAGACACACAGTTGTTCCCCACAGCGCCAGATGTG 5125
QY 170 GGACCCCTCTCTCCCGCAGATTTCTCTTCATATCTTCTTTATCCATCCG 229
Db 5124 GAAGTACTCACTCTCTCCGAGCTGCTTCCTTCATGAGCCCTGACCTGCTCC 5065
QY 230 TTGAAGCAACGCACTA 246
Db 5064 TGGTAGCAGCTGTACCA 5048
RESULT 13
ABV04394
ID ABV04394 standard; cDNA; 386 BP.
XX
AC ABV04394;
XX
DT 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 4385.
DE
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200108660-A2.
XX
PD 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JR;
XX WPI, 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 761, 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or incidence of prostate cancer in a patient
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SO Sequence 386 BP; 108 A; 74 C; 74 G; 128 T; 0 U; 2 Other;
Query Match 5.2%; Score 37.4; DB 5; Length 386;
Best Local Similarity 64.4%; Pred. No. 0.95;
Matches 56; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 628 ACTATGACAGTTCTGGACGCTTAACCTTTATGTCAGAGACTATCAATCATACA 687
DB 175 ATATGTGGAACCTCACTGCTATTTTCAGGTTCTCTTAAGTACTAAGGAACTCCA 234
QY 688 GATATTTGTCAAAAAAGACTAA 714
DB 235 GATACCTTCAAAAAAAGCTAA 261
RESULT 14
AAL35477/c
ID AAL35477 standard; cDNA; 366 BP.
XX AAL35477;
AC 08-JAN-2002 (first entry)
XX
DT Human musculoskeletal system related polynucleotide SEQ ID NO 819.
XX
DE
XX
KW Cytostatic; immunosuppressive; nocitropic; neuroprotective; antiviral;
KW antiinfective; hepatotropic; antidiabetic; antiinflammatory; anticancer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ss.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001338.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0225799P.
PR 22-AUG-2000; 2000US-0226811P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.

20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX
XX MPI: 2001-451937/48.
XX P-PSDB; ABB03895.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
XX PT disorders related to the musculoskeletal system including musculoskeletal
XX PT cancers and also for testing and detection e.g. diagnosis.
PS Claim 1; SEQ ID NO 819; 781bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX CC (ABB03087-ABB04109) associated with the musculoskeletal system useful for
XX CC preventing, treating or ameliorating medical conditions e.g. by protein
XX CC or gene therapy. The genes are isolated from a range of human tissues
XX CC disclosed in the specification. The nucleic acids, proteins, antibodies

CC and (ant)agonists are useful in the diagnosis, treatment and prevention
CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 366 BP; 120 A; 59 C; 53 G; 133 T; 0 U; 1 Other;
SQ
Query Match 5.1%; Score 37; DB 4; Length 366;
Best Local Similarity 50.6%; Pred. No. 1.2;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 549 AGAGATTAGCAAGCGGAAAAAACTATGCTGAGCTGAGCTTTTTCATCATATATA 608
DB 205 AGATTATTATTTTAAATGAATTAATAAATGAGCGCTTTTAAATGCTTAATTTTAAA 146
QY 609 AAGGAGAAATGTTGCTCCTCATATGTACACTTTCTGGAGCTCTTAATTTTATTCAG 668
DB 145 GAAAGACCAAGTGTCTTCTCTATGCTGCAACCAATTCGAATGTTGAGTTAAGTAA 86
QY 669 AGCACTATCAATCATCATAGATATGTGCAAAAAAAAAGACATATATATACA 722
DB 85 AACTTTTCAATTAANACAGTCATTATACGTTAAAAAAAATCCAGTTGATA 32
RESULT 15
ABXS8465/C
ID ABXS8465 standard; CDNA; 366 BP.
AC
XX ABXS8465;
XX
XX
DT 26-FEB-2003 (first entry)
XX
XX
DE CDNA encoding novel human musculoskeletal system antigen #809.
XX
XX
KW Gene; ss; musculoskeletal system antigen; cancer; metastasis;
KW re-vascularisation; thrombosis; arteriosclerosis; mineral content;
KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
KW post-operative tissue repair; limb regeneration; neuronal growth;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW AIDS-related complex; chondrocyte growth; bone regeneration;
KW periodontal regeneration; tissue transport; bone graft; skin aging;
KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KW cell growth; organ transplant; cell differentiation; body height; weight;
KW hair colour; eye colour; skin; percentage of adipose tissue;
KW pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
KW depression; tendency for violence; pain; reproductive capability;
KW hormone level; endocrine level; appetite; libido; memory; stress;
KW storage capability; fat content; lipid content; protein content;
KW carbohydrate content; vitamin content; cofactor content;
KW nutritional component.
OS Homo sapiens.
XX
XX US2002147140-A1.
XX
XX 10-OCT-2002.
XX
XX
XX 17-JAN-2001; 2001US-00764877.
XX
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 28-JUN-2000; 2000US-0214886P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.

keratinocyte growth; prevents hair loss; since FGF family members activate hair-forming cells and promotes melanocyte growth; stimulates growth and differentiation of hematopoietic cells and bone marrow cells when used in combination with other cytokines; maintains organs before transplantation or for supporting cell culture of primary tissues; induces tissue of mesodermal origin to differentiate in early embryonic increases or decreases the differentiation or proliferation of embryonic stem cells; besides, hematopoietic lineage; modulates mammalian characteristics, such as, body height, weight, hair colour, eye colour, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery); modulates mammalian metabolism; changes mammal's metal state or physical state by influencing biohythms, circadian rhythms, depression, tendency for violence, tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, or stress; increases or decreases storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components. This sequence encodes a novel human musculoskeletal system antigen. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at
ftp.segdata.uspto.gov/sequence.html?DocID=20020147140

SQ Sequence 366 BP; 120 A; 59 C; 53 G; 133 T; 0 U; 1 Other;

Query Match 5.1%; Score 37; DB 8; Length 366;
Best Local Similarity 50.6%; Pred. No. 1.2;
Matches 89; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

OY AGAGAAATTAGCAACGCGAAAAAACAATATGGCTAGCTGGAGTTGTTTTCAATCATATA 608
DB 205 AGATTATTTTTTAAATGAATAAATAAATGCGCCTTTTAATAGTTGTTAATTTGCAAA 146
OY 609 AAGGAGAAATTGTGCTCACTATGTACAGTTTCGSGACGCTTAACTTTAATTCAG 668
DB 145 GAAGGACCAAGTGTTCCTCTCTATGCTGCAACCAATTCCAAATGTTGGCTTTTAGGTA 86
OY 669 AGGACTATCAATATATACAGATATTGCCAAAAAAGACTAATAATATACA 722
DB 85 AACCTTTCAATAAANAAGTCATTATACCCTTAAAAAATCCAGATTGATA 32

Search completed: August 27, 2005, 08:07:49
Job time : 534 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 09:08:05 ; Search time 656 Seconds
(without alignments)
7211.363 Million cell updates/sec

Title: US-10-776-213-2

Sequence: 1 cttcgattagcagcagcacac.....aaaagactaataataacat 723

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	723	100.0	723	19	US-10-776-213-2
2	723	100.0	11437	19	US-10-776-213-24
3	723	100.0	13073	19	US-10-776-213-20
4	720.4	99.6	850	19	US-10-776-213-30
5	497.4	68.8	680	19	US-10-451-467A-317
c	41	5.7	1024	14	US-10-123-155-198
c	7	5.7	1024	15	US-10-146-731-198

c	8	41	5.7	1024	15	US-10-140-472-198	Sequence 198, App
c	9	41	5.7	1024	15	US-10-141-761-198	Sequence 198, App
c	10	41	5.7	1024	15	US-10-142-885-198	Sequence 198, App
c	11	41	5.7	1024	16	US-10-158-790-198	Sequence 198, App
c	12	41	5.7	1024	17	US-10-137-871-198	Sequence 198, App
c	13	41	5.7	1024	17	US-10-140-923-198	Sequence 198, App
c	14	41	5.7	1024	17	US-10-141-756-198	Sequence 198, App
c	15	41	5.7	1024	17	US-10-141-759-198	Sequence 198, App
c	16	41	5.7	1024	17	US-10-140-805-198	Sequence 198, App
c	17	41	5.7	1024	17	US-10-140-864-198	Sequence 198, App
c	18	41	5.7	1024	18	US-10-142-446-198	Sequence 198, App
c	19	41	5.7	2561	17	US-10-197-824-20	Sequence 20, App1
c	20	39.4	5.4	409	18	US-10-424-599-136845	Sequence 23, App1
c	21	38.6	5.3	2311	19	US-10-466-531-23	Sequence 175, App
c	22	38.4	5.3	544	14	US-10-066-543-1775	Sequence 277, App
c	23	38.4	5.3	1225	21	US-10-956-157-277	Sequence 278, App
c	24	38.4	5.3	1225	21	US-10-956-157-278	Sequence 5512, App
c	25	38.4	5.3	1225	21	US-10-956-157-5512	Sequence 5513, App
c	26	38.4	5.3	1225	21	US-10-956-157-5513	Sequence 29921, App
c	27	38.4	5.3	2355	18	US-10-424-599-29921	Sequence 129944
c	28	38.2	5.3	1635	18	US-10-424-599-129944	Sequence 333, App
c	29	37.8	5.2	7736	9	US-09-764-846-333	Sequence 333, App
c	30	37.8	5.2	7736	14	US-10-091-483-333	Sequence 4385, App
c	31	37.4	5.2	386	20	US-10-357-930-4385	Sequence 96559, App
c	32	37.4	5.2	436	18	US-10-424-599-96559	Sequence 106894
c	33	37.2	5.1	685	13	US-10-027-632-106894	Sequence 106895
c	34	37.2	5.1	685	13	US-10-027-632-106895	Sequence 106894
c	35	37.2	5.1	685	17	US-10-027-632-106894	Sequence 106895
c	36	37.2	5.1	685	17	US-10-027-632-106895	Sequence 25182, App
c	37	37.2	5.1	704	13	US-10-027-632-25182	Sequence 25183, App
c	38	37.2	5.1	704	17	US-10-027-632-25182	Sequence 94399, App
c	39	37.2	5.1	704	17	US-10-027-632-25182	Sequence 819, App
c	40	37.2	5.1	1121	18	US-10-424-599-94399	Sequence 819, App
c	41	37.2	5.1	366	9	US-09-764-877-819	Sequence 1976, App
c	42	37	5.1	366	17	US-09-764-877-819	Sequence 2883, App
c	43	37	5.1	366	17	US-09-960-352-1976	
c	44	37	5.1	366	17	US-09-960-352-1976	
c	45	37	5.1	400	11	US-09-969-034-2883	

ALIGNMENTS

RESULT 1
US-10-776-213-2
; Sequence 2, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: Astrazeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/10/776, 213
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-776-213-2

Query Match 100.0%; Score 723; DB 19; Length 723;
Best Local Similarity 100.0%; Pred. No. 1.1e-216;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CTTCGATTAGCAGCAGACATCATGAGCTGCTATATAATACATCGAAAA	60
Db	1	CTTCGATTAGCAGCAGACATCATGAGCTGCTATATAATACATCGAAAA	60
Qy	61	ACCATTAAGAGCAAGCGATACCTTCTGGAAGAAAGAGCAGCGCTTTGAAGGGGAT	120
Db	61	ACCATTAAGAGCAAGCGATACCTTCTGGAAGAAAGAGCAGCGCTTTGAAGGGGAT	120

```
QY 121 GGGGGCTAAGAAATCATTCATCTTCTTCCCTTCGGGTCGGAAACCCGGGACCCCTCTT 180
DB 121 GGGGGCTAAGAAATCATTCATCTTCTTCCCTTCGGGTCGGAAACCCGGGACCCCTCTT 180
QY 181 CTCCCGGCAAGATTTCTTCTTCAATCTCTCTTTTATTTCTCATCTCCGTTGAAGAAC 240
DB 181 CTCCCGGCAAGATTTCTTCTTCAATCTCTCTTTTATTTCTCATCTCCGTTGAAGAAC 240
QY 241 GCATATGACTAATATGATGATGATCTCAATGAGCTGATGATGATGATGATGATGATGAT 300
DB 241 GCATATGACTAATATGATGATGATCTCAATGAGCTGATGATGATGATGATGATGATGAT 300
QY 301 GGTAAAGGACCGTGGCTCGAAACGGTTCTTCTGACAAATTCAGAACAGGGGCTACA 360
DB 301 GGTAAAGGACCGTGGCTCGAAACGGTTCTTCTGACAAATTCAGAACAGGGGCTACA 360
QY 361 GTCTGATATATGAAATATATAGGCAATTTTGTATAGCGCGCGCGCGCGCGCTTTCCA 420
DB 361 GTCTGATATATGAAATATATAGGCAATTTTGTATAGCGCGCGCGCGCGCGCTTTCCA 420
QY 421 ATAGGAGGCGGAGTTATCGGCGAGCTCTACTTCTTCTTGGGTAAGCCCTTTC 480
DB 421 ATAGGAGGCGGAGTTATCGGCGAGCTCTACTTCTTCTTGGGTAAGCCCTTTC 480
QY 481 TGTCTTTCGCGCAAGTGTGCTGCGAGCTGCGCGAGAACATATGATATAGGAGTATAC 540
DB 481 TGTCTTTCGCGCAAGTGTGCTGCGAGCTGCGCGAGAACATATGATATAGGAGTATAC 540
QY 541 TTTTCGATGAGAAATATGAGAGGGAATAATATATGCTAGCTGGAGTTGTTTCAA 600
DB 541 TTTTCGATGAGAAATATGAGAGGGAATAATATATGCTAGCTGGAGTTGTTTCAA 600
QY 601 TCATATATAAGGAGAAATTTGCTCAGTATGACAGTTTCTGGGAGCTTAACTTT 660
DB 601 TCATATATAAGGAGAAATTTGCTCAGTATGACAGTTTCTGGGAGCTTAACTTT 660
QY 661 TATTGCAAGAGCATATCAATCATACAGATATTTGCAAAAAAAGACTATATATA 720
DB 661 TATTGCAAGAGCATATCAATCATACAGATATTTGCAAAAAAAGACTATATATA 720
QY 721 CAT 723
DB 721 CAT 723
```

RESULT 2
US-10-776-213-24

```
; Sequence 24, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: AstraZeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/10/776,213
; CURRENT FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 11427
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-776-213-24
```

Query Match 100.0%; Score 723; DB 19; Length 11427;
Best Local Similarity 100.0%; Pred. No. 4,7e-216;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CTTTGATTAGCAGCAGACATCATATGATGCTGCTATTAATAATACACTAGGAGAAA 60
DB 15 CTTTGATTAGCAGCAGACATCATATGATGCTGCTATTAATAATACACTAGGAGAAA 74
```

```
QY 61 ACCATMAAGACGAATACCTTGAAGAGAAAGAGACGCTTTGAAAGGGAGAT 120
DB 75 ACCATMAAGACGAATACCTTGAAGAGAAAGAGACGCTTTGAAAGGGAGAT 134
QY 121 GGGGGCTAAGAAATCATTCATCTTCCCTTCGGGTCGGAAACCCGGGACCCCTCTT 180
DB 121 GGGGGCTAAGAAATCATTCATCTTCCCTTCGGGTCGGAAACCCGGGACCCCTCTT 180
QY 181 CTCCCGGCAAGATTTCTTCTTCAATCTCTCTTTTATTTCTCATCTCCGTTGAAGAAC 240
DB 181 CTCCCGGCAAGATTTCTTCTTCAATCTCTCTTTTATTTCTCATCTCCGTTGAAGAAC 240
QY 241 GCATATGACTAATATGATGATGATCTCAATGAGCTGATGATGATGATGATGATGATGAT 300
DB 241 GCATATGACTAATATGATGATGATCTCAATGAGCTGATGATGATGATGATGATGATGAT 300
QY 301 GGTAAAGGACCGTGGCTCGAAACGGTTCTTCTGACAAATTCAGAACAGGGGCTACA 360
DB 301 GGTAAAGGACCGTGGCTCGAAACGGTTCTTCTGACAAATTCAGAACAGGGGCTACA 360
QY 361 GTCTGATATATGAAATATATAGGCAATTTTGTATAGCGCGCGCGCGCGCGCTTTCCA 420
DB 361 GTCTGATATATGAAATATATAGGCAATTTTGTATAGCGCGCGCGCGCGCGCTTTCCA 420
QY 421 ATAGGAGGCGGAGTTATCGGCGAGCTCTACTTCTTCTTGGGTAAGCCCTTTC 480
DB 421 ATAGGAGGCGGAGTTATCGGCGAGCTCTACTTCTTCTTGGGTAAGCCCTTTC 480
QY 481 TGTCTTTCGCGCAAGTGTGCTGCGAGCTGCGCGAGAACATATGATATAGGAGTATAC 540
DB 481 TGTCTTTCGCGCAAGTGTGCTGCGAGCTGCGCGAGAACATATGATATAGGAGTATAC 540
QY 541 TTTTCGATGAGAAATTTAGCAGGGAATAATATATGCTAGCTGGAGTTGTTTCAA 600
DB 541 TTTTCGATGAGAAATTTAGCAGGGAATAATATATGCTAGCTGGAGTTGTTTCAA 600
QY 601 TCATATATAAGGAGAAATTTGCTCAGTATGACAGTTTCTGGGAGCTTAACTTT 660
DB 601 TCATATATAAGGAGAAATTTGCTCAGTATGACAGTTTCTGGGAGCTTAACTTT 660
QY 661 TATTGCAAGAGCATATCAATCATACAGATATTTGCAAAAAAAGACTATATATA 720
DB 661 TATTGCAAGAGCATATCAATCATACAGATATTTGCAAAAAAAGACTATATATA 720
QY 721 CAT 723
DB 721 CAT 723
QY 735 CAT 737
DB 735 CAT 737
```

RESULT 3
US-10-776-213-20

```
; Sequence 20, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: AstraZeneca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/10/776,213
; CURRENT FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 13073
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-776-213-20
```

Query Match 100.0%; Score 723; DB 19; Length 13073;
Best Local Similarity 100.0%; Pred. No. 5.1e-216;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CTTTGATTAGCAGCAGACATCATATGATGCTGCTATTAATAATACACTAGGAGAAA 60
```

```
|||||
Db 16 CTTTCATTAGCGACACACATCACTAGACTGCGTCAATAAATACACTACGAGAAA 75
Qy 61 ACCATAAGGCAAGCAAGTACTACTTGGAAAGAAAGAGACCCCTTGTAAGGGGAT 120
Db 76 ACCATAAAGGCAAGCAAGTACTACTTGGAAAGAAAGAGACCCCTTGTAAGGGGAT 135
Qy 121 GGGGGCTAAGAAATCACTTCTTTTCCCTTCCGCGGTCCGGACCCGAGACCCCTCT 180
Db 136 GGGGGCTAAGAAATCACTTCTTTTCCCTTCCGCGGTCCGGACCCGAGACCCCTCT 195
Qy 181 CTCGCCGACAGATTTCTTCTTCAATCTTCTTTTATTCCTATCCGTTGAAGCAAC 240
Db 196 CTCGCCGACAGATTTCTTCTTCAATCTTCTTTTATTCCTATCCGTTGAAGCAAC 255
Qy 241 GCACTATGACTAATGCTGTGACATCTTCAGATGCTGTGACTGTGTATCTCAAGT 300
Db 256 GCACTATGACTAATGCTGTGACATCTTCAGATGCTGTGACTGTGTATCTCAAGT 315
Qy 301 GGTAAAGGCAAGCTGCTCGAAACGGTTCCTTCTGTACAAATCTAGAAACAGGGCTACA 360
Db 316 GGTAAAGGCAAGCTGCTCGAAACGGTTCCTTCTGTACAAATCTAGAAACAGGGCTACA 375
Qy 361 GTCCTGATATAGAAATTAATAGGCAATTTTGTCTAGCGCGCGCGCGCTTTCCCA 420
Db 376 GTCCTGATATAGAAATTAATAGGCAATTTTGTCTAGCGCGCGCGCGCTTTCCCA 435
Qy 421 ATAGGAGGCGGAGTTTATGCGGAGCTTACTTTCTTATTTGGGTAAAGCCCTTTC 480
Db 436 ATAGGAGGCGGAGTTTATGCGGAGCTTACTTTCTTATTTGGGTAAAGCCCTTTC 495
Qy 481 TGTTCGCGGCAAGTGTGCTGAGGCTGCGCGGAGCAATAGATTAAGGATGTAAAC 540
Db 496 TGTTCGCGGCAAGTGTGCTGAGGCTGCGCGGAGCAATAGATTAAGGATGTAAAC 555
Qy 541 TTTGATGAGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 600
Db 556 TTTGATGAGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 615
Qy 601 TCATATTAAGGAGAAATTTGCTGCTCATATGTGACATTTTGGAGCTTAACTTT 660
Db 616 TCATATTAAGGAGAAATTTGCTGCTCATATGTGACATTTTGGAGCTTAACTTT 675
Qy 661 TATTCGAGAGCACTATCAATCATACAGATTTGTCAAAAAAAGACTAATATAA 720
Db 676 TATTCGAGAGCACTATCAATCATACAGATTTGTCAAAAAAAGACTAATATAA 735
Qy 721 CAT 723
Db 736 CAT 738

RESULT 4
US-10-776-213-30
; Sequence 30, Application US/10776213
; Publication No. US20040142478A1
; GENERAL INFORMATION:
; APPLICANT: Astrazenca AB
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for Controlling Nucleic Acid Expression
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/10/776,213
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-776-213-30

Query Match 99.6%; Score 720.4; DB 19; Length 850;
Best Local Similarity 99.9%; Pred. No. 7,6e-216;
```

```
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CTTTCATTAGCGACACACATCACTAGACTGCGTCAATAAATACACTACGAGAAA 60
Db 78 CTTTCATTAGCGACACACATCACTAGACTGCGTCAATAAATACACTACGAGAAA 137
Qy 61 ACCATAAAGGCAAGCAAGTACTACTTGGAAAGAAAGAGACCCCTTGTAAGGGGAT 120
Db 138 ACCATAAAGGCAAGCAAGTACTACTTGGAAAGAAAGAGACCCCTTGTAAGGGGAT 197
Qy 121 GGGGGCTAAGAAATCACTTCTTTTCCCTTCCGCGGTCCGGACCCGAGACCCCTCT 180
Db 198 GGGGGCTAAGAAATCACTTCTTTTCCCTTCCGCGGTCCGGACCCGAGACCCCTCT 257
Qy 181 CTCGCCGACAGATTTCTTCTTCAATCTTCTTTTATTCCTATCCGTTGAAGCAAC 240
Db 258 CTCGCCGACAGATTTCTTCTTCAATCTTCTTTTATTCCTATCCGTTGAAGCAAC 317
Qy 241 GCACTATGACTAATGCTGTGACATCTTCAGATGCTGTGACTGTGTATCTCAAGT 300
Db 318 GCACTATGACTAATGCTGTGACATCTTCAGATGCTGTGACTGTGTATCTCAAGT 377
Qy 301 GGTAAAGGCAAGCTGCTCGAAACGGTTCCTTCTGTACAAATCTAGAAACAGGGCTACA 360
Db 378 GGTAAAGGCAAGCTGCTCGAAACGGTTCCTTCTGTACAAATCTAGAAACAGGGCTACA 437
Qy 361 GTCCTGATATAGAAATTAATAGGCAATTTTGTCTAGCGCGCGCGCGCTTTCCCA 420
Db 438 GTCCTGATATAGAAATTAATAGGCAATTTTGTCTAGCGCGCGCGCGCTTTCCCA 497
Qy 421 ATAGGAGGCGGAGTTTATGCGGAGCTTACTTTCTTATTTGGGTAAAGCCCTTTC 480
Db 498 ATAGGAGGCGGAGTTTATGCGGAGCTTACTTTCTTATTTGGGTAAAGCCCTTTC 557
Qy 481 TGTTCGCGGCAAGTGTGCTGAGGCTGCGCGGAGCAATAGATTAAGGATGTAAAC 540
Db 558 TGTTCGCGGCAAGTGTGCTGAGGCTGCGCGGAGCAATAGATTAAGGATGTAAAC 617
Qy 541 TTTGATGAGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 600
Db 618 TTTGATGAGAGAAATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 677
Qy 601 TCATATTAAGGAGAAATTTGCTGCTCATATGTGACATTTTGGAGCTTAACTTT 660
Db 678 TCATATTAAGGAGAAATTTGCTGCTCATATGTGACATTTTGGAGCTTAACTTT 737
Qy 661 TATTCGAGAGCACTATCAATCATACAGATTTGTCAAAAAAAGACTAATATAA 720
Db 738 TATTCGAGAGCACTATCAATCATACAGATTTGTCAAAAAAAGACTAATATAA 797
Qy 721 CA 722
Db 798 AA 799

RESULT 5
US-10-451-467A-317
; Sequence 317, Application US/10451467A
; Publication No. US20040161840A1
; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN YEAST AND FUNGI
; FILE REFERENCE: JAB-1667
; CURRENT APPLICATION NUMBER: US/10/451,467A
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: EP 00870318.3
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: EP 01870002.1
; PRIOR FILING DATE: 2001-01-04
```

```
; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 317
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-451-467A-317

Query Match      68.8%; Score 497.4; DB 19; Length 680;
Best Local Similarity 99.8%; Pred. No. 1.2e-145;
Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 224 ATCCCGTTGAAGCAACCGCACTATGACTAATGCTGAGACATCTCCATGGCTGACT 283
DB 1 ATCCCGTTGAAGCAACCGCACTATGACTAATGCTGAGACATCTCCATGGCTGACT 60
QY 284 TGTGTATCTCACTGAGTGAACGGCAACCGCTGCGAAACGGTTCTTCTGTGACATTT 343
DB 61 TGTGTATCTCACTGAGTGAACGGCAACCGCTGCGAAACGGTTCTTCTGTGACATTT 120
QY 344 CTAGAACAGGGGCTACAGTCTCGATATAGATATAGAGGCAATTTTGTGAGCGCGCC 403
DB 121 CTAGAACAGGGGCTACAGTCTCGATATAGATATAGAGGCAATTTTGTGAGCGCGCC 180
QY 404 GCGGCGCCGCTTCCCAATAGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTAT 463
DB 181 GCGGCGCCGCTTCCCAATAGAGGCGCAGTTTATCGGCGAGCTCTACTTCTTCTAT 240
QY 464 TTGGGTAACCCCTTCTGTTTTCGGCCAGTGTCTCTGAGGCTGCGCGGAGAACATA 523
DB 241 TTGGGTAACCCCTTCTGTTTTCGGCCAGTGTCTCTGAGGCTGCGCGGAGAACATA 300
QY 524 GTGATAGAGGATGATCTTTCGATGAGAGATTTAGCAGCGGAAAAAATCTATGGCTAGC 583
DB 301 GTGATAGAGGATGATCTTTCGATGAGAGATTTAGCAGCGGAAAAAATCTATGGCTAGC 360
QY 584 TGGGAGTTGTTTTCATCATATATAAGGAGAAATTTGCTCACTATGTACAGTTTC 643
DB 361 TGGGAGTTGTTTTCATCATATATAAGGAGAAATTTGCTCACTATGTACAGTTTC 420
QY 644 TGGGAGTTGTTTTCATCATATATAAGGAGAAATTTGCTCACTATGTACAGTTTC 703
DB 421 TGGGAGTTGTTTTCATCATATATAAGGAGAAATTTGCTCACTATGTACAGTTTC 480
QY 704 AAAAAGACTATATATATACA 722
DB 481 AAAAAGACTATATATATAAA 499

RESULT 6
US-10-123-155-198/c
; Sequence 198, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
```

```
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; PRIOR FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-198

Query Match      5.7%; Score 41; DB 14; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.1;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

QY 47 TACACTAGGAAAAACCTAAAGAGCAAGGATCTACTTGGAGAAAAAGAGCAGC 106
DB 954 T.CRYAN..MMMHV..MWT..HYCD..NYTBD..BH..TH..H..BB..SN..S..N..YN..NYHS 895
QY 107 CTYTAAGGGGATGGGCTAAGAAGCATTCATCTTTCCTTCCCTTCCGCGTCCGAGC 166
DB 894 CA.T...MM..SBTASD..MKW..WBH..H..SBCAGRB..HYBN...YNYSS..WHS..A 835
QY 167 CCGGACCCCTCCCTCTCCCGCAGATTTCTTCTTCATATCTTCTTTATTCCTATC 226
DB 834 ST..DGDNMC..SGT..K..B..YV...D..RST..B..SCN..NM..WC..M..RAATB..M..R..N..N 775
QY 227 CCGTTGAAGCAACCGCACTATGACTAATGAGTGTCTGAGCATCTGCATGCTGTGACTGT 286
DB 774 RN..KTH..MH..R..TYSSTDW...HM..S..RY.....N...RCYLT..S..THH...CTYNS 715
QY 287 GTGATCTCACAGTGTGAACGCGACCGTGGCTCGGAAACGGTCTCTGTGACATTTCTA 346
DB 714 BS..B..DNY...H...YMWY..HRNBX..RCA..N..NC..WSCM..RA..YD..SWMSB..T..S 655
QY 347 GAAACAGGGCTACAGTCTCGATATATAGATATAGCGCATTTTGTAGCGCGCGCG 406
DB 654 SBDSYNCB..A...W..RSNN..M..TMAMTS..HR..D..A...YN..TANC..A..B..RCK 595
QY 407 GCGCCCGTTCCCAATAGAGGCGCGAGTTTATGCGGAGCTACTTCTTCTATTTG 466
DB 594 ..AM..HXH...NBRCNVM..T..S..ANW..HNYTTCRMD..A..RH..CY..BBDCK..NT..KY 535
QY 467 GTTAGCCCTTTTCTGTTTGGCGCAGTGTGCTGCGAGCTGCGCGCGAGAACATAGTG 526
DB 534 ..TTCMRT..WH..BYH..HAT..B...SB..DSST..CH..KB..D..BHA..Y..M..KWC..R..RH 475
QY 527 ATAAAGATGTACTTTCGATGAGAGATTTAGCAAGCGGAAAAAATATAGGCTAGCTG 586
DB 474 ACB...RDT..C..M..WBH..MWB..AB..HC..W..DG...DB..BKABH..RS..SSBS..H...Y..B 415
QY 587 GAGTTGTTTTCATCATATATAAGGAGAAATTTGCTCACTATGTGACAGTTTCGCG 646
DB 414 ..S..CWN..KTR..DM..BT..H..M...SR..BM..RH..R...YVA..D..TRCSCYV..TH..YTR 355
QY 647 GAGCTTAATCTTATATGCGAGAGCATATCAATTCATACATACATATATGTCAAAAAAA 706
DB 354 .YBBSKMAAA..YAA..NG..TWT..RRASK..A..S..DHABWCH...BSA..B..KNS..S..PAT. 295
QY 707 AAGACTATATATATAC 721
DB 294 DM..S..M..ATMR...H 280

RESULT 7
US-10-146-731-198/c
; Sequence 198, Application US/10146731
; Publication No. US20030129692A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
```

```

/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaeroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K
/ APPLICANT: Zhang, Zemin
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3330R1C323
/ CURRENT APPLICATION NUMBER: US/10/146,731
/ CURRENT FILING DATE: 2002-05-15
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 198
/ LENGTH: 1024
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-146-731-198

```

Query Match 5.7%; Score 41; DB 15; Length 1024;

Best Local Similarity 7.7%; Pred. No. 0.1;

Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

```

QY 47 TACACTACGGAAGAAACATTAAGAGCAAGCGATACCTCTTGGAAGAAAGAGCAGC 106
DB 954 T.CRYAN..MMHY..MMT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYHS 895
QY 107 CTGTGAAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 166
DB 894 GA.T...MM.SBTAASD.MKW..WBH.H..SBGAGR.BHYBN...YVNS.WHS..A 835
QY 167 CCGGAGCCCTCTCTCTCCCGCAGATTTCTTTCATATCTTCTTATCTTCTTATC 226
DB 834 ST..DEDMNC.SGT.K.B.YY..D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CGGTGAAGCAACCGCATATGCTAAATGCTGAGCATCTCCATGCTGAGTGT 286
DB 774 RN.KYH..MH..R.TYSTDW..HM..S.RY.....N..RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTCAGAGTGTAACGCGACCGGCTCGAAGCGTTCTTCTGACATTTCTA 346
DB 714 BS.B.DNY..H..YVNY.HRNBV.RCA.N.NC..WSCMH.RA.YDD..SMNSBW..T.S 655
QY 347 GAACAGGGGCTACGCTCGATATAGAAATATATAGCGCATTTTCTAGCGCGCGC 406
DB 654 SBDSTNCB.A...W.RSNN..M.TMAMTS..HR..D..A..YN.TAANC..A.B.RCK 595
QY 407 GCGCCGTTTCCCAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 466
DB 594 ..AM.HKB...NBRMCNYM.T.S.ANW..HNYTNGMD.A.RH..CY..BSPDK..NT..KY 535
QY 467 GGTAAAGCCCTTCTCTTCTGCGCAGTGCTGCTGAGGCTGCGCGGAGGAGCATAGTG 526
DB 534 ..TTCMRT.MH.BYH.HAT..B..SB.DSST..CH..KB..D.BHA..Y.M.KWIC.R.RH 475
QY 527 ATTAAGGATGTAACTTTCATGTAGAGAAATTAAGAGCGGAAAAAACTATGCTAGCTGG 586
DB 474 ACB...RDT.C.M.WBH.WMB..AB..HC.W.DG..DB.BKABH..RS.SBS..H...Y..B 415
QY 587 GAGTGTGTTTTCATATATAAAGGAGAAATTTGTCCTCATATGTGACATTTCTGG 646
DB 414 ..S.CMN.KTTR..DM..BT..H.M..SR.BM.RH.R...YMA..D..TRHSCYV..TH..YTR 355
QY 647 GAGCTTAATCTTTATGCGAGAGCATATCAATACATACAGATATTTGCAAAAAAAA 706

```

```

DB 354 .YBBSCMAAA.YRA.NG.TWT.RRASK..A.S.DHAEBCH...BSA.B.KUS.S.RAT. 295
QY 707 AAGACTAATATATAC 721
DB 294 DM.S.M.ATWR...H 280

```

RESULT 8

US-10-140-472-198/C

/ Sequence 198, Application US/10140472

/ Publication No. US2003013888A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Beresini, Maureen

/ APPLICANT: DeForge, Laura

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Filvaeroff, Ellen

/ APPLICANT: Gao, Wei-Qiang

/ APPLICANT: Gerritsen, Mary E.

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Sherwood, Steven

/ APPLICANT: Smith, Victoria

/ APPLICANT: Stewart, Timothy A.

/ APPLICANT: Tumas, Daniel

/ APPLICANT: Watanabe, Colin K

/ APPLICANT: Zhang, Zemin

/ APPLICANT: Wood, William

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

/ FILE REFERENCE: P3330R1C168

/ CURRENT APPLICATION NUMBER: US/10/140,472

/ CURRENT FILING DATE: 2002-05-06

/ Prior Application removed - See File Wrapper or Palm

/ NUMBER OF SEQ ID NOS: 550

/ SEQ ID NO 198

/ LENGTH: 1024

/ TYPE: PRT

/ ORGANISM: Homo Sapien

US-10-140-472-198

Query Match 5.7%; Score 41; DB 15; Length 1024;

Best Local Similarity 7.7%; Pred. No. 0.1;

Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

```

QY 47 TACACTACGGAAGAAACATTAAGAGCAAGCATACCTTGGAAGAAAGAGCAGC 106
DB 954 T.CRYAN..MMHY..MMT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYHS 895
QY 107 CTGTGAAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 166
DB 894 GA.T...MM.SBTAASD.MKW..WBH.H..SBGAGR.BHYBN...YVNS.WHS..A 835
QY 167 CCGGAGCCCTCTCTCTCCCGCAGATTTCTTTCATATCTTCTTATCTTCTTATC 226
DB 834 ST..DEDMNC.SGT.K.B.YY..D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CGGTGAAGCAACCGCATATGCTAAATGCTGAGCATCTCCATGCTGAGTGT 286
DB 774 RN.KYH..MH..R.TYSTDW..HM..S.RY.....N..RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTCAGAGTGTAACGCGACCGGCTCGAAGCGTTCTTCTGACATTTCTA 346
DB 714 BS.B.DNY..H..YVNY.HRNBV.RCA.N.NC..WSCMH.RA.YDD..SMNSBW..T.S 655
QY 347 GAACAGGGGCTACGCTCGATATAGAAATATATAGCGCATTTTCTAGCGCGCGC 406
DB 654 SBDSTNCB.A...W.RSNN..M.TMAMTS..HR..D..A..YN.TAANC..A.B.RCK 595
QY 407 GCGCCGTTTCCCAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 466
DB 594 ..AM.HKB...NBRMCNYM.T.S.ANW..HNYTNGMD.A.RH..CY..BSPDK..NT..KY 535
QY 467 GGTAAAGCCCTTCTCTTCTGCGCAGTGCTGCTGAGGCTGCGCGGAGGAGCATAGTG 526
DB 534 ..TTCMRT.MH.BYH.HAT..B..SB.DSST..CH..KB..D.BHA..Y.M.KWIC.R.RH 475
QY 527 ATTAAGGATGTAACTTTCATGTAGAGAAATTAAGAGCGGAAAAAACTATGCTAGCTGG 586
DB 474 ACB...RDT.C.M.WBH.WMB..AB..HC.W.DG..DB.BKABH..RS.SBS..H...Y..B 415
QY 587 GAGTGTGTTTTCATATATAAAGGAGAAATTTGTCCTCATATGTGACATTTCTGG 646
DB 414 ..S.CMN.KTTR..DM..BT..H.M..SR.BM.RH.R...YMA..D..TRHSCYV..TH..YTR 355
QY 647 GAGCTTAATCTTTATGCGAGAGCATATCAATACATACAGATATTTGCAAAAAAAA 706

```

```

Db      594 ..AM.HKB...NBRMCNVM.T.S.ANV.HNYTTNCRMD.A.RH..CY..BSDCK.NT.KY.535
QY      467 GGTAAAGCCCTTTCTGTTTCGCGCAGTGGTTCGTCAGAGCGCGCGGAGAACATAGTG.526
Db      534 ..TTTCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH.475
QY      527 ATAAAGGATGTAACTTTCGATGAGAGATTGCAAGCGGAAAAAACTATGGCTAGCTGG.586
Db      474 ACB...RDT.C.M.WBH.WMB.AB.HC.W.DG..DB.BKABH..RS.SBSB.H...Y..B.415
QY      587 GAGTGTCTTTTCATCATATATAAAGGAGAAATTTGCTCACTATGTACAGTTTCTGG.646
Db      414 ..S.CMN.KTRR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHCSGY.YH.YTR.355
QY      647 GACGCTTAACTTTATGTCAGAGACTATCAATCATCATATATTGCAAAAAAAA.706
Db      354 .YBBBCSMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT.295
QY      707 AAGACTAATAATAAC.721
Db      294 DM.S.M.ATMR...H.280

```

RESULT 9

```

US-10-141-761-198/c
; Sequence 198, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-761-198

```

Query Match 5.7%; Score 41; DB 15; Length 1024;

```

Best Local Similarity 7.7%; Pred. No. 0.1;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;
QY      47 TCACACGAGGAAAAACATTAAGAGCAAGCGACTACTCTTGGAGAGAAAGAGACAG.106
Db      954 T.CRYAN..MMWY..MMT.HYCD.NYTB.D..BH.TH.H.BB.SN.S.N..YN.NYHHS.895
QY      107 CTTGTAAGGGGATGAGGGGCTAGAGATCATCTTTCTTCCCTGCGGCTCGGAC.166
Db      894 GA.T...MM.SPTASD.MKW...WBH.H..SBGAGR.BYBN...YVYNS.WHS..A.835
QY      167 CCGGACCCCTCTCTCCCGGACGAGATTTCTTCTTCATATCTTCTTTATCTTCTTANC.226
Db      834 ST..DGDWNC.SGT.X.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N.775

```

```

QY      227 CCGTTGAAGCAACCGCACTATGACTAATGCTGCGACATCTCCAGCTGTGACTTGT.286
Db      774 RN.KYH..MH..R.TYSTTDW...HM..S.RY.....N...RCTYT.S.THH..CTVNS.715
QY      287 GTGTATTCACAGTGTAAAGGACCGCGCTGCTCGGAAAACGTTCTTGTGTGACATTTCTA.346
Db      714 BS.B.DNY...H...YMYV..HRNBX.RCA.N.NC..WSCMH.RA.YD.D.SMNSBW..T.S.655
QY      347 GAACAGGGGCTACAGTCTCGATTAATAGATTAAGCCATTTTGTCTAGCGCCGCCGCG.406
Db      654 SBDSINCB.A...W.RSNM..M.TMAMTS.HR..D..A..YN.TAANC..A.B.RCK.595
QY      407 GCGCCGTTTCCCATAGAGAGCGCAGTTTATGCGGAGCTTACTTCTTCTATTGT.466
Db      594 ..AM.HKB...NBRMCNVM.T.S.ANV.HNYTTNCRMD.A.RH..CY..BSDCK.NT.KY.535
QY      467 GGTAAAGCCCTTTCTGTTTCGCGCAGTGGTTCGTCAGAGCGCTGCGCGCGGAGAACATAGTG.526
Db      534 ..TTTCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH.475
QY      527 ATAAAGGATGTAACTTTCGATGAGAGATTGCAAGCGGAAAAAACTATGGCTAGCTGG.586
Db      474 ACB...RDT.C.M.WBH.WMB.AB.HC.W.DG..DB.BKABH..RS.SBSB.H...Y..B.415
QY      587 GAGTGTCTTTTCATCATATATAAAGGAGAAATTTGCTCACTATGTGACAGTTTCTGG.646
Db      414 ..S.CMN.KTRR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHCSGY.YH.YTR.355
QY      647 GACGCTTAACTTTATGTCAGAGACTATCAATCATCATATATTGCAAAAAAAA.706
Db      354 .YBBBCSMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT.295
QY      707 AAGACTAATAATAAC.721
Db      294 DM.S.M.ATMR...H.280

```

RESULT 10

```

US-10-142-885-198/c
; Sequence 198, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C248
; CURRENT APPLICATION NUMBER: US/10/142,885
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-885-198

```

```

Query Match      5.7%; Score 41; DB 16; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.1;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

QY 47 TACACTACGAAAAACCATTAAGACCAAGCCTACTCTTGAAGAAAAAGACACG 106
DB 954 T.CRYAN..MMMHY..MT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYNHS 895
QY 107 CTGTGAAGGGGAGGAGCTAAGATCATCTTCTTTCCCTGCGCGCGAC 166
DB 894 GA.T...MM.SBTAASD..MKW..WBH.H..SBGAGR.B.HYBN...YVNS.S.WHS..A 835
QY 167 CCGGAGACCCCTCTCTCCCGGACGATTTCTTCTTCAATCTCTTTATCTCTATC 226
DB 834 ST..DGDNMC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CCGTTGAAGCAACCGCATATGCTTAATGCTGCTGACATCTCCATGCTGACTTGT 286
DB 774 RN.KYH..MH..R.TYSTTDM...HM..S.RY.....N...RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTCAGATGTGAACGACGCGTGCAGAAAGCTTCTGACATTTCTA 346
DB 714 BS.B.DNY...H...YVNY.HRNB.Y.RCA.N.NC..WSCMH.RA.YDD.SMNSBW..T.S 655
QY 347 GAACAGGGGCTACAGTCTCGATATGAATTAATTAAGCGCATTTTCTAGCGCGCGG 406
DB 654 SBDSYNCBB.A...W.RSNN..M.TMAMTS.HR..D..A..YN.TAANC..A.B.RCK 595
QY 407 GCGCCCGCTTCCCAATAGGAGGCGCATTTATGCGCGGAGCTCTACTTCTTCTATTTG 466
DB 594 ..AM.HKB...NBRMCNMY.T.S.ANW..HNYTNCMD.A.RH..CY..BSDCK.NT.KY 535
QY 467 GGTAAAGCCCTTCTGTTTTCGCGCAGTGTGCTGACGCGCGCGGAGCAATAGTG 526
DB 534 ..TTCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KWNC.R.RH 475
QY 527 ATAAAGGATGTAACTTTCGATGAGAGATTAAGCAACGCGAAAAAATAATGCTAGCTG 586
DB 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B 415
QY 587 GAGTTGTTTCAATCATATATAAAGGAGAAATGTTGCTCACTATGTGACAGTTTCTG 646
DB 414 ..S.CMN.KTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYV.TH.YTR 355
QY 647 GAGCTTTAATCTTTATGAGAGAGACTATCAATCATACAGATATGTCAAAAAAAA 706
DB 354 ..YBBBSMAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATAAC 721
DB 294 DM.S.M.ATMR...H 280

RESULT 11
US-10-158-790-198/c
; Sequence 198, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tuma, Daniel
; APPLICANT: Watanabe, Colin K

```

```

; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C448
; CURRENT APPLICATION NUMBER: US/10/158,790
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-158-790-198

Query Match      5.7%; Score 41; DB 16; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.1;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;

QY 47 TACACTACGAAAAACCATTAAGACCAAGCCTACTCTTGAAGAAAAAGACACG 106
DB 954 T.CRYAN..MMMHY..MT.HYCD..NYTBD..BH.TH.H.BB..SN.S.N..YN..NYNHS 895
QY 107 CTGTGAAGGGGAGGAGCTAAGATCATCTTCTTTCCCTGCGCGCGAC 166
DB 894 GA.T...MM.SBTAASD..MKW..WBH.H..SBGAGR.B.HYBN...YVNS.S.WHS..A 835
QY 167 CCGGAGACCCCTCTCTCCCGGACGATTTCTTCTTCAATCTCTTTATCTCTATC 226
DB 834 ST..DGDNMC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
QY 227 CCGTTGAAGCAACCGCATATGCTTAATGCTGCTGACATCTCCATGCTGACTTGT 286
DB 774 RN.KYH..MH..R.TYSTTDM...HM..S.RY.....N...RCTYT.S.THH..CTVNS 715
QY 287 GTGTATCTCAGATGTGAACGACGCGTGCAGAAAGCTTCTGACATTTCTA 346
DB 714 BS.B.DNY...H...YVNY.HRNB.Y.RCA.N.NC..WSCMH.RA.YDD.SMNSBW..T.S 655
QY 347 GAACAGGGGCTACAGTCTCGATATGAATTAATTAAGCGCATTTTCTAGCGCGCGG 406
DB 654 SBDSYNCBB.A...W.RSNN..M.TMAMTS.HR..D..A..YN.TAANC..A.B.RCK 595
QY 407 GCGCCCGCTTCCCAATAGGAGGCGCATTTATGCGCGGAGCTCTACTTCTTCTATTTG 466
DB 594 ..AM.HKB...NBRMCNMY.T.S.ANW..HNYTNCMD.A.RH..CY..BSDCK.NT.KY 535
QY 467 GGTAAAGCCCTTCTGTTTTCGCGCAGTGTGCTGACGCGCGCGGAGCAATAGTG 526
DB 534 ..TTCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KWNC.R.RH 475
QY 527 ATAAAGGATGTAACTTTCGATGAGAGATTAAGCAACGCGAAAAAATAATGCTAGCTG 586
DB 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B 415
QY 587 GAGTTGTTTCAATCATATATAAAGGAGAAATGTTGCTCACTATGTGACAGTTTCTG 646
DB 414 ..S.CMN.KTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYV.TH.YTR 355
QY 647 GAGCTTTAATCTTTATGAGAGAGACTATCAATCATACAGATATGTCAAAAAAAA 706
DB 354 ..YBBBSMAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KNS.S.RAT. 295
QY 707 AAGACTAATATAAC 721
DB 294 DM.S.M.ATMR...H 280

RESULT 12
US-10-137-871-198/c
; Sequence 198, Application US/10137871
; Publication No. US20030207350A1
; GENERAL INFORMATION:

```



```

QY 407 GCGCCGTTTCCCAATAGGAGCGAGATTATCGCGGAGCTTACTTCTTCTATTG 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 594 ..AM.HKB...NBRMCNYM.T.S.ANW..HNYTTCMD.A.RH..CY..BSDCK.NT.KY 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 GGTAAAGCCCTTCTGTTTTCGGCCAGTGTGCTGCGAGGCTGCCGGAACAATAGT 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 534 ..TTCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH 475
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 527 ATAGAGGATGTACTTTCATGAGAGTAATGCAAGCGGAAAAAACTATGGCTACTG 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 587 GAGTGTCTTTCATCATATTAAGGAGAAATGTGCTCACTATGTGACAGTTTCTG 646
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 ..S.CMN.KTTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYV.TH.YTR 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 647 GAGCTCTTAACCTTTATTCGAGAGACTATCAATCATACAGATATTGTCAAAAAA 706
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 ..YBBBCSMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KXS.S.RAT. 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 707 AAGACTAATATATAC 721
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 DM.S.M.ATMR...H 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 14

```

US-10-141-756-198/c
; Sequence 198, Application US/10141756
; Publication No. US20030207359A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C200
; CURRENT APPLICATION NUMBER: US/10/141,756
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-756-198

```

```

Query Match 5.7%; Score 41; DB 17; Length 1024;
Best Local Similarity 7.7%; Pred. No. 0.1;
Matches 52; Conservative 194; Mismatches 429; Indels 0; Gaps 0;
QY 47 TACACTACGAAAAACCATTAAGAGCAAGGATACCTTGTGAAGAAAGAGACAG 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 954 T.CRYAN..MMHY..MWT.HYCD.NYTB..BH.TH.H.BB.SN.S.N..YN..NYHS 895
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 107 CTGTAAAGGAGATGGGCTTAAGAGATCATTTCTTTCCCTCGCGCTCCGAG 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 894 GA.T...NM.SBTASD.MKW...WBH.H..SBGARB.HYBN...YNNNS.WHS..A 835
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 167 CCGGAGCCCTCTCTCCCGGACGATTTCTTCAATATCTCTTTATTCGATC 226
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 834 ST..DGDNMC.SGT.K.B.YY...D.RST.B.SCN.NNM.WC.M.RAATB.M..R.N..N 775
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 227 CCGTTGAAGCAACCGACATATGACTTAATGCTGCTGAGACATCTCCATGCTGACTTGT 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 774 RN.KYH..MH..R.TYSTTDW...HM..S.RY.....N..RCTYT.S.THH..CTYNS 715
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 287 GTGTATCTCAGAGTGTGTAACGACCGCTGCTCGGAAACGTTCTTCTGTAATTTCA 346
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 714 BS.B.DNY...H..YMY.HRNBV.RCA.N.NC..WSCMH.RA.YDD.SMBSBM..T.S 655
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 347 GAACAGGGCTACAGCTCGATTAATGAATTAATTAACCGCATTTTGTAGCGCCGCG 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 654 SBDSYNCB.A...W.RSNN..M.TMAVTS.HR..D..A...YN.TAANC..A.B.RCK 595
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 407 GCGCCGTTTCCCAATAGGAGCGAGCTTATCGCGGAGCTTACTTCTTCTATTG 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 594 ..AM.HKB...NBRMCNYM.T.S.ANW..HNYTTCMD.A.RH..CY..BSDCK.NT.KY 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 GGTAAAGCCCTTCTGTTTTCGGCCAGTGTGCTGCGAGGCTGCCGGAACAATAGT 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 534 ..TTCMRT.WH.BYH.HAT..B..SB.DSST.CH..KB.D.BHA..Y.M.KNWC.R.RH 475
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 527 ATAGAGGATGTACTTTCATGAGAGTAATGCAAGCGGAAAAAACTATGCTACTG 586
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 474 ACB...RDT.C.M.WBH.WWB.AB.HC.W.DG..DB.BKABH..RS.SBS.H...Y..B 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 587 GAGTGTCTTTCATCATATTAAGGAGAAATGTGCTCACTATGTGACAGTTTCTG 646
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 ..S.CMN.KTTR.DM.BT..H.M..SR.BM.RH.R...YMA.D.TRHSCYV.TH.YTR 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 647 GAGCTCTTAACCTTTATTCGAGAGACTATCAATCATACAGATATTGTCAAAAAA 706
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 ..YBBBCSMAAA.YRA.NG.TWT.RRASK..A.S.DHABWCH...BSA.B.KXS.S.RAT. 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 707 AAGACTAATATATAC 721
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 DM.S.M.ATMR...H 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

```

US-10-141-759-198/c
; Sequence 198, Application US/10141759
; Publication No. US20030207361A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C197
; CURRENT APPLICATION NUMBER: US/10/141,759
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 198
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Homo Sapien

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2005, 07:32:20 ; Search time 178 Seconds
(without alignments)
6646.225 Million cell updates/sec

Title: US-10-776-213-2
Perfect score: 723
Sequence: 1 cttcgattagcagcagcac.....aaaagacataataacat 723

Scoring table: IDENTITY_NOC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	723	100.0	723	US-09-743-194-2	Sequence 2, Appl1
2	723	100.0	11427	US-09-743-194-24	Sequence 24, Appl1
3	723	100.0	13073	US-09-743-194-20	Sequence 20, Appl1
4	720.4	99.6	850	US-09-743-194-30	Sequence 30, Appl1
5	47	6.5	7218	US-08-232-463-14	Sequence 14, Appl1
6	37.8	5.2	4989	US-09-949-016-17445	Sequence 17445, A
7	37.8	5.2	11440	US-09-949-016-17442	Sequence 17442, A
8	36.6	5.1	16662	US-09-949-016-16850	Sequence 16850, A
9	35.8	5.0	815	US-09-369-247-15	Sequence 15, Appl1
10	35.4	4.9	191433	US-09-949-016-16144	Sequence 16144, A
11	35.4	4.9	212449	US-09-949-016-15419	Sequence 15419, A
12	35.2	4.9	72992	US-09-949-016-17532	Sequence 17532, A
13	34.8	4.8	98962	US-09-949-016-14133	Sequence 14133, A
14	34.8	4.8	102884	US-09-949-016-17100	Sequence 17100, A
15	34.4	4.8	601	US-09-949-016-26160	Sequence 26160, A
16	34.4	4.8	601	US-09-949-016-18345	Sequence 18345, A
17	34.4	4.8	51723	US-09-949-016-12152	Sequence 12152, A
18	34.4	4.8	51723	US-09-949-016-16931	Sequence 16931, A
19	34.4	4.8	105189	US-09-949-016-13029	Sequence 13029, A
20	34.2	4.7	1876	US-09-469-242-3	Sequence 3, Appl1
21	34.2	4.7	14205	US-09-949-016-16196	Sequence 16196, A
22	34.2	4.7	42571	US-09-810-347-3	Sequence 3, Appl1
23	34	4.7	168394	US-09-949-016-13631	Sequence 13631, A
24	34	4.7	168394	US-09-949-016-13002	Sequence 13002, A
25	33.8	4.7	232547	US-09-949-016-16603	Sequence 16603, A
26	33.8	4.7	313	US-09-513-999C-35864	Sequence 35864, A
27	33.8	4.7	502	US-09-621-976-18891	Sequence 18891, A

28	33.8	4.7	4818	US-08-817-926-27	Sequence 27, Appl1
29	33.6	4.6	77851	US-09-949-016-12508	Sequence 12508, A
30	33.6	4.6	77851	US-09-949-016-13211	Sequence 13211, A
31	33.6	4.6	77867	US-09-949-016-13212	Sequence 13212, A
32	33.6	4.6	77940	US-09-949-016-12509	Sequence 12509, A
33	33.4	4.6	2172	US-09-976-594-125	Sequence 125, App
34	33.4	4.6	312470	US-09-949-016-14043	Sequence 14043, A
35	33.4	4.6	336024	US-09-949-016-12373	Sequence 12373, A
36	33.2	4.6	784019	US-09-949-016-14033	Sequence 14033, A
37	33.2	4.6	828152	US-09-949-016-12777	Sequence 12777, A
38	33	4.6	601	US-09-949-016-28524	Sequence 28524, A
39	33	4.6	601	US-09-949-016-60957	Sequence 60957, A
40	33	4.6	205163	US-09-949-016-17009	Sequence 17009, A
41	32.8	4.5	62776	US-09-949-016-17576	Sequence 17576, A
42	32.8	4.5	98567	US-09-949-016-11750	Sequence 11750, A
43	32.8	4.5	100567	US-09-949-016-16934	Sequence 16934, A
44	32.8	4.5	301828	US-09-949-016-13969	Sequence 13969, A
45	32.6	4.5	243	US-09-248-796A-8463	Sequence 8463, Ap

ALIGNMENTS

RESULT 1					
US-09-743-194-2					
; Sequence 2, Application US/09743194					
; Patent No. 6716601					
; GENERAL INFORMATION:					
; APPLICANT: Bellfield, Graham					
; APPLICANT: Oakley, Caroline					
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for					
; FILE REFERENCE: 3526.82543					
; CURRENT APPLICATION NUMBER: US/09/743,194					
; CURRENT FILING DATE: 2001-01-08					
; NUMBER OF SEQ ID NOS: 32					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 2					
; LENGTH: 723					
; TYPE: DNA					
; ORGANISM: Saccharomyces cerevisiae					
; US-09-743-194-2					
Query Match 100.0%; Score 723; DB 4; Length 723;					
Best Local Similarity 100.0%; Pred. No. 1,9e+229; Indels 0; Gaps 0;					
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	CTTTCGATTAGCAGCAGACATCATAGACTGCTCATATAAATACATACGAAAA	60		
DB	1	CTTTCGATTAGCAGCAGACATCATAGACTGCTCATATAAATACATACGAAAA	60		
QY	61	ACCATTAAGAGCAAGCATCTACTTTGGAGAAAAAGAGCAGCTTTGAAGGGGAT	120		
DB	61	ACCATTAAGAGCAAGCATCTACTTTGGAGAAAAAGAGCAGCTTTGAAGGGGAT	120		
QY	121	GGGGGCTAAGAGTCACTTCTTTCCCTTCGCGGTCCGAGCCCGGAGCCCTCT	180		
DB	121	GGGGGCTAAGAGTCACTTCTTTCCCTTCGCGGTCCGAGCCCGGAGCCCTCT	180		
QY	181	CTTCCCGCAGATTTCTTCTTCATATCTCTTTATCTATCCGTTGAAGAAC	240		
DB	181	CTTCCCGCAGATTTCTTCTTCATATCTCTTTATCTATCCGTTGAAGAAC	240		
QY	241	GCATCTAGCTAATGTGTCTGACATCTTCATGTGCTGATCTGTATCTACAG	300		
DB	241	GCATCTAGCTAATGTGTCTGACATCTTCATGTGCTGATCTGTATCTACAG	300		
QY	301	GGTAAAGGACCGTGGCTCGAAAAAGCTTCTTGTGACAAATCTGAACAGGGCTACA	360		
DB	301	GGTAAAGGACCGTGGCTCGAAAAAGCTTCTTGTGACAAATCTGAACAGGGCTACA	360		
QY	361	GTCTCATATATAGATTAATTAAGGCAATTTTGTCTACGCGCGCGCGCGCTTCCCA	420		

```
Db 361 GTCTCGATTAATAGAAATATAGACCAATTTTGTCTAGCGCCGCGCGCCGCTTTCCCA 420
Qy 421 ATAGGAGGCGCAGTTTATCGCGGAGCTCTACTTCTTCTTATTTGGTAAAGCCCTTTTC 480
Db 421 ATAGGAGGCGCAGTTTATCGCGGAGCTCTACTTCTTCTTATTTGGTAAAGCCCTTTTC 480
Qy 481 TGTTCCTGCGCAATGTTGTCTGAGGCTGCGCGGAAACATGATGATTAAGGATGTAAAC 540
Db 481 TGTTCCTGCGCAATGTTGTCTGAGGCTGCGCGGAAACATGATGATTAAGGATGTAAAC 540
Qy 541 TTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATGAGCTGAGGAGTGTGTTTTTCAA 600
Db 541 TTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATGAGCTGAGGAGTGTGTTTTTCAA 600
Qy 601 TCATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTTCTGGGACGCTTTAACTTT 660
Db 601 TCATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTTCTGGGACGCTTTAACTTT 660
Qy 661 TATTGCAAGAGACTATCAAAATCTATCAGATATTGTCAAAAAAAAGACTATATATA 720
Db 661 TATTGCAAGAGACTATCAAAATCTATCAGATATTGTCAAAAAAAAGACTATATATA 720
Qy 721 CAT 723
Db 721 CAT 723

RESULT 2
US-09-743-194-24
; Sequence 24, Application US/09743194
; Patent No. 6716601
; GENERAL INFORMATION:
; APPLICANT: Bellfield, Graham
; APPLICANT: Oakley, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 11427
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-743-194-24

Query Match 100.0%; Score 723; DB 4; Length 11427;
Best Local Similarity 100.0%; Pred. No. 9.8e-229; Indels 0; Gaps 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTCGATTAGCAGCAGACATCATAGACTGCGTCATATAAAATACACTACGGAANA 60
Db 15 CTTTCGATTAGCAGCAGACATCATAGACTGCGTCATATAAAATACACTACGGAANA 74
Qy 61 ACCATTAAGAGCAAAAGCATCTTCTTGAAGAAAAAGAGCAGCGTTGTAAGGGGGAT 120
Db 75 ACCATTAAGAGCAAAAGCATCTTCTTGAAGAAAAAGAGCAGCGTTGTAAGGGGGAT 134
Qy 121 GGGGGCTAAGAGTATTCATCTTCTTCCCTTCCGCGGTCCGGAGCCCGGGACCCCTCT 180
Db 135 GGGGGCTAAGAGTATTCATCTTCTTCCCTTCCGCGGTCCGGAGCCCGGGACCCCTCT 194
Qy 181 CTCCTCCGACAGATTTCTTCTTCAATCTTCTTATTTCTATCCCGGTGAAGCAAC 240
Db 195 CTCCTCCGACAGATTTCTTCTTCAATCTTCTTATTTCTATCCCGGTGAAGCAAC 254
Qy 241 GCACATATGACTAATATGCTGCTGAGACATCTCATGCGCTGACTTGTGTATCTCAGT 300
Db 255 GCACATATGACTAATATGCTGCTGAGACATCTCATGCGCTGACTTGTGTATCTCAGT 314
Qy 301 GGTAAAGGACCGGTGCTCGGAAAGGTTCCCTTGTGACAAATTTCTAGAAAGGGGCTACA 360
Db 361 GGTAAAGGACCGGTGCTCGGAAAGGTTCCCTTGTGACAAATTTCTAGAAAGGGGCTACA 360
```

```
Db 315 GGTAAAGGACCGGTGCTCGGAAAGGTTCCCTTGTGACAAATTTCTAGAAAGGGGCTACA 374
Qy 361 GTCTCGATTAATAGAAATATAGACCAATTTTGTCTAGCGCCGCGCGCCGCTTTCCCA 420
Db 375 GTCTCGATTAATAGAAATATAGACCAATTTTGTCTAGCGCCGCGCGCCGCTTTCCCA 434
Qy 421 ATAGGAGGCGCAGTTTATCGCGGAGCTCTACTTCTTCTTATTTGGTAAAGCCCTTTTC 480
Db 435 ATAGGAGGCGCAGTTTATCGCGGAGCTCTACTTCTTCTTATTTGGTAAAGCCCTTTTC 494
Qy 481 TGTTCCTGCGCAATGTTGTCTGAGGCTGCGCGGAAACATGATGATTAAGGATGTAAAC 540
Db 495 TGTTCCTGCGCAATGTTGTCTGAGGCTGCGCGGAAACATGATGATTAAGGATGTAAAC 554
Qy 541 TTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATGAGCTGAGGAGTGTGTTTTTCAA 600
Db 555 TTTCGATGAGAAATTTAGCAAGCGGAAAAAACTATGAGCTGAGGAGTGTGTTTTTCAA 614
Qy 601 TCATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTTCTGGGACGCTTTAACTTT 660
Db 615 TCATATTAAGGAGAAATTTGTCTCACTATGTGACAGTTTCTGGGACGCTTTAACTTT 674
Qy 661 TATTGCAAGAGACTATCAAAATCTATCAGATATTGTCAAAAAAAAGACTATATATA 720
Db 675 TATTGCAAGAGACTATCAAAATCTATCAGATATTGTCAAAAAAAAGACTATATATA 734
Qy 721 CAT 723
Db 735 CAT 737

RESULT 3
US-09-743-194-20
; Sequence 20, Application US/09743194
; Patent No. 6716601
; GENERAL INFORMATION:
; APPLICANT: Bellfield, Graham
; APPLICANT: Oakley, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 13073
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-743-194-20

Query Match 100.0%; Score 723; DB 4; Length 13073;
Best Local Similarity 100.0%; Pred. No. 1.1e-228; Indels 0; Gaps 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTTCGATTAGCAGCAGACATCATAGACTGCGTCATATAAAATACACTACGGAANA 60
Db 16 CTTTCGATTAGCAGCAGACATCATAGACTGCGTCATATAAAATACACTACGGAANA 75
Qy 61 ACCATTAAGAGCAAAAGCATCTTCTTGAAGAAAAAGAGCAGCGTTGTAAGGGGGAT 120
Db 76 ACCATTAAGAGCAAAAGCATCTTCTTGAAGAAAAAGAGCAGCGTTGTAAGGGGGAT 135
Qy 121 GGGGGCTAAGAGTATTCATCTTCTTCCCTTCCGCGGTCCGGAGCCCGGGACCCCTCT 180
Db 136 GGGGGCTAAGAGTATTCATCTTCTTCCCTTCCGCGGTCCGGAGCCCGGGACCCCTCT 195
Qy 181 CTCCTCCGACAGATTTCTTCTTCAATCTTCTTATTTCTATCCCGGTGAAGCAAC 240
Db 196 CTCCTCCGACAGATTTCTTCTTCAATCTTCTTATTTCTATCCCGGTGAAGCAAC 255
Qy 241 GCACATATGACTAATATGCTGCTGAGACATCTCATGCGCTGACTTGTGTATCTCAGT 300
Db 255 GCACATATGACTAATATGCTGCTGAGACATCTCATGCGCTGACTTGTGTATCTCAGT 300
```

Db 256 GCACATATGACTAATAGTGTGTCGACATCTCCATAGCTGTGATCTGTGTATCTCAAGT 315
QY 301 GGTAACGGACCGTGGCTCGGAAACGGTTCTTCTGTACAAATCTGAACAGGGGCTACA 360
Db 316 GGTAACGGACCGTGGCTCGGAAACGGTTCTTCTGTACAAATCTGAACAGGGGCTACA 375
QY 361 GTCTCGATTAATGAATTAATGAACGGCAATTTTGTCTAGCGCCGCGCGCCGCTTTCCCA 420
Db 376 GTCTCGATTAATGAATTAATGAACGGCAATTTTGTCTAGCGCCGCGCGCCGCTTTCCCA 435
QY 421 ATAGGAGGCGGAGTTTATTCGGGAGCTCTACTTCTTCTTATTTGGGTAAGCCCTTTTC 480
Db 436 ATAGGAGGCGGAGTTTATTCGGGAGCTCTACTTCTTCTTATTTGGGTAAGCCCTTTTC 495
QY 481 TGTTCGCGCCAGTGTGTCTGACAGGCTGCGCGGAGAACATAGTATTAAGGATGTAAAC 540
Db 496 TGTTCGCGCCAGTGTGTCTGACAGGCTGCGCGGAGAACATAGTATTAAGGATGTAAAC 555
QY 541 TTTGATGAGAGAAATTTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTTCAA 600
Db 556 TTTGATGAGAGAAATTTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTTCAA 615
QY 601 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 660
Db 616 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 675
QY 661 TATTGAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATATA 720
Db 676 TATTGAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATATA 735
QY 721 CAT 723
Db 736 CAT 738

RESULT 4
US-09-743-194-30
; Sequence 30, Application US/09743194
; Patent No. 6716601
; GENERAL INFORMATION:
; APPLICANT: Bellfield, Graham
; APPLICANT: Oakley, Caroline
; TITLE OF INVENTION: Compositions and Methods Utilizing Sequences for
; FILE REFERENCE: 3526.82543
; CURRENT APPLICATION NUMBER: US/09/743,194
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 30
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; US-09-743-194-30

Query Match 99.6%; Score 720.4; DB 4; Length 850;
Best Local Similarity 99.9%; Pred. No. 1.5e-228;
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTCGATTAGCAGCAGACATCATAGACTGCGTCATATAAAATACACTACGAAAA 60
Db 78 CTTTCGATTAGCAGCAGACATCATAGACTGCGTCATATAAAATACACTACGAAAA 137
QY 61 ACCATAAAGCAAGCAAGCACTACTTGTGAGGAAAAAGAGACCGCTTGTAAAGGGGAT 120
Db 138 ACCATAAAGCAAGCAAGCACTACTTGTGAGGAAAAAGAGACCGCTTGTAAAGGGGAT 197
QY 121 GGGGGCTAAGAAATCACTTCTTTCCCTGCGGATCGGAGCCCGGACCCCTCC 180
Db 198 GGGGGCTAAGAAATCACTTCTTTCCCTGCGGATCGGAGCCCGGACCCCTCC 257
QY 181 CTCGCCGACGATTTCTCTTCAATCTCTCTTTATCTCTAATCCGTTGAAGCAAC 240

Db 258 CTCGCCGACGATTTCTTCTTCAATATCTCTTTTATCTCTAATCCGTTGAAGCAAC 317
QY 241 GCACATATGACTAATAGTGTGTCGACATCTCCATAGCTGTGATCTGATCTCAAGT 300
Db 318 GCACATATGACTAATAGTGTGTCGACATCTCCATAGCTGTGATCTGATCTCAAGT 377
QY 301 GGTAACGGACCGTGGCTCGGAAACGGTTCTTCTGTACAAATCTGAACAGGGGCTACA 360
Db 376 GGTAACGGACCGTGGCTCGGAAACGGTTCTTCTGTACAAATCTGAACAGGGGCTACA 437
QY 361 GTCTCGATTAATGAATTAATGAACGGCAATTTTGTCTAGCGCCGCGCGCCGCTTTCCCA 420
Db 438 GTCTCGATTAATGAATTAATGAACGGCAATTTTGTCTAGCGCCGCGCGCCGCTTTCCCA 497
QY 421 ATAGGAGGCGGAGTTTATTCGGGAGCTCTACTTCTTCTTATTTGGGTAAGCCCTTTTC 480
Db 498 ATAGGAGGCGGAGTTTATTCGGGAGCTCTACTTCTTCTTATTTGGGTAAGCCCTTTTC 557
QY 481 TGTTCGCGCCAGTGTGTCTGACAGGCTGCGCGGAGAACATAGTATTAAGGATGTAAAC 540
Db 558 TGTTCGCGCCAGTGTGTCTGACAGGCTGCGCGGAGAACATAGTATTAAGGATGTAAAC 617
QY 541 TTTGATGAGAGAAATTTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTTCAA 600
Db 618 TTTGATGAGAGAAATTTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTTCAA 677
QY 601 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 660
Db 678 TCATATTAAGGAGAAATTTGTTGCTCACTATGTGACAGTTTCTGGAGCTCTTAACCTT 737
QY 661 TATTGAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATATA 720
Db 738 TATTGAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATATA 797
QY 721 CA 722
Db 798 AA 799

RESULT 5
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768

```

:
: REFERENCE/DOCKET NUMBER: 30472/114 IMM
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: (703) 836-9300
:
: TELEFAX: (703) 683-4109
:
: TELEX: 899149
:
: INFORMATION FOR SEQ ID NO: 14:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 7218 base pairs
:
: TYPE: nucleic acid
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
: IMMEDIATE SOURCE:
:
: CLONE: pTZpT-F15
:
: US-08-232-463-14

```

Query Match 6.5%; Score 47; DB 1; Length 7218;
Best Local Similarity 2.9%; Pred. No. 0.00024;
Matches 11; Conservative 212; Mismatches 152; Indels 0; Gaps 0;

QY	113	AGGGGGAATGGGGGCTAAGAAGTCATTCACTTCTTTCCCTCGCGGCAGACCCGGGA	172
Db	1047	AGGTTCAGAGGAGACTTGCGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1106
QY	173	CCCCCTCCTCCCAGCAGATTCTTCTCTCATATCTTCTTTATTCCTATGCCGTG	232
Db	1107	TT	1166
QY	233	AAGCACCCGACTATGACTAAATGCTGCTGACATCTTCATGCGCTGACTTGCTGAT	292
Db	1167	TT	1226
QY	293	CTCACAAGTGAACGGCACCCTGGCTGGAAAGCGTTCCTCGTAGCAATCTAGAACAG	352
Db	1227	TT	1286
QY	353	GGGCTACAGTCTCGATAATAGATAATAGCGCAATTTTGTACGCCGCCGCGCC	412
Db	1287	TT	1346
QY	413	GTTTCCCAATAGGAGGGCGCAGTTATTCGCGGAGCTCTACTTCTCTAATTTGGGTAG	472
Db	1347	TT	1406
QY	473	CCCCTTCTGTTTC	487
Db	1407	TTTTTTTTTTTTTTTT 1421	

```

RESULT 6
US-09-949-016-17445
; Sequence 17445, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17445
; LENGTH: 4989
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17445

```

Query Match	5.2%;	Score 37.8;	DB 4;	Length 4989;
Best Local Similarity	54.7%;	Pred. No. 0.22;		
Matches 75;	Conservative 0;	Mismatches 62;	Indels 0;	Gaps 0;

Qy	110	GTAAAGGGGAGTGGGGGTAGGAAGTCACTTCTTTTCCCTTCGCGTTCGGAGACCG	169
Db	4464	GTTTCAGGGGATGGATGTATTAAGCACACACAGTTTCTCCCCCACAGCCGCCAGATGTG	4523
Qy	170	GGACCCCTCCTCTCCCGCACAGATTCTCTTCAATCTCTTATTTCTATCCCG	229
Db	4524	GAAGTACTTCACCTCTCTCCCGAGTGTGCTTTCCTTCATAGCCTCTGACCTCCCTCCC	4583
Qy	230	TTGAAGCAACGCACCTA	246
Db	4584	TGGTAGCAGCTGTACCA	4600

```

RESULT 7
US-09-949-016-17442/c
; Sequence 17442, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001107
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,438
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 17442
; LENGTH: 11440
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17442

```

```

Query Match: 5.2%; Score 37.8; DB 4; Length 11440;
Best Local Similarity 54.7%; Pred. No. 0.36; Indels 0; Gaps 0;
Matches 75; Conservative 0; Mismatches 62;

QY      110 GTAAGGGGGATGGGGGCTAAGAGTCATCACTTCTCTTCCCTTGCGGGTCCGACCCG 169
Db      8773 GTTCAAGGGGATGGATGTAAAGCACAACACAGTGTCTTCCGCCACAGCCCGAGATGTG 8714

QY      170 GGACCCCTCTCTCTCCCGCAGATTTCTCTCTTCATATCTCTTTATTCATCCG 229
Db      8713 GAAATATTCACCTCTCCCTCCGAGTGTGCTTTCCCTCATGGCCCTGACCTGCTCC 8654

QY      230 TTGAAGCAACCGCACTA 246
Db      8653 TGTGACAGCTGTACCA 8637

RESULT 8
US-09-949-016-16850/c
; Sequence 16850, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

```

? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 607/231,498
? PRIOR FILING DATE: 2000-03-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FASTSEQ for Windows Version 4.0.0
? SEQ ID NO 16850
? LENGTH: 16662
? TYPE: DNA
? ORGANISM: Human
? MS-09-949-016-16850

```

Query Match	5.1%	Score 36.6	DB 4	Length 16662
Best Local Similarity	54.0%	Pred. No.	1.1	
Matches 75	Conservative	0	Mismatches 64	Indels 0
				Gaps 0

QY	89	GGAGGAAAAGAGACACCCCTGTAAAGGGGATGGGGGCTAAGAAAGTCAATTCATCTTCTT	148
Db	4397	GGAGGAATATGACATATCATATATAAAGGCGTCAGAGAGTGGCTTTTCTTCTCCTT	4338
QY	149	TTCCTTGGGGTCCGGACCCGGGACCCCTCTCTCCCGGACGATTTCTTCTCTTCAAT	208
Db	4337	CTCCTTCTTCTTCTTCTCTCTCTCTCTCTGCAACACACACTCTCTCTCTTCTCT	4278

Qy	209	CTTCCTTTATTCCTATCC	227
Db	4277	CTTCCTTTCTTTTCTTTCC	4259

RESULT 5 US-09-369-247-15
Sequence 15, Application US/09369247
Patent No. 656992
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 44 Human Secreted Proteins
FILE REFERENCE: P2024p1
CURRENT APPLICATION NUMBER: US/09/369, 247
CURRENT FILING DATE: 1999-08-05
EARLIER APPLICATION NUMBER: 60/074,118
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,157
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,137
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,341
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,141
EARLIER FILING DATE: 1998-02-09
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PatentIn Ver. 2.0

```

? LENGTH: 815
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (406)
? OTHER INFORMATION: n equals a,t,g, or c
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (794)
? OTHER INFORMATION: n equals a,t,g, or c
? IS-09-369-247-15

```

Query Match	5.0%	Score	35.8	DB	4	Length	815
Best Local Similarity	55.1%	Pred. No.	0.34				
Matches	70	Conservative	0	Mismatches	57	Indels	0
				Gaps			0

Oy 594 TTTTCATCATATAAAGGAGAATTGTGCTCACTGTGCACGTTTCCGACGCT 655
 | | | | | | | | | | | | | | | |
Db 665 TATCCTGTATTAAATGTGAACATTATTTGACACTTCAGTGAGTTATAGCTTAATAGT 724
 | | | | | | | | | | | | | | | |
Oy 654 TAACTTTTATTGCGAGAGCATCAAATATACAGATATTGTCAAAAAAAAAAAGACTA 713

D_b
725 CTTGCGCTATGCAGCAGGTGTAATAATTAAATAATTTTTTAAATAAAAAAAAAAAAA 784

Qy
714 ATATATNA 720

D_b
785 AAAAAAAAA 791

RESULT 10
US-09-949-016-16144
; Sequence 16144, Application US/09949016
; Patent No. 6812339

```

? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: C10011307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,458
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for windows Version 4.0
? SEQ ID NO 16144
? LENGTH: 19143
? TYPE: DNA
? ORGANISM: Human
? IS-09-949-016-16144

```

Query Match Similarity	4.98;	Score 35.4;	DB 4;	Length 191433;
Best Local Similarity	59.48;	Pred. No.12;		
Matches 60; Conservative	0;	Mismatches 41;	Indels 0;	Gaps 0;

QY 104 CCGCTTGTAGGGGAGTGGGGGCTAGAAACATTCCTTTCCCTTGGGGTCCG 163

Db 41249 ACAGTGGTAAAGGAGAGAGTACATCAGTGTTCACAGTCTTTTTTCTCCAGTAAG 41308

Oy 164 GACCCGGGACCCTCCTCTTCCCCGCAGATTTCCTTC 204
||| ||||| ||||| |||||
Db 41309 TCCCCCCCACCCCCCGGCTGCCCTTGCCGATTTTC 41349

RESULT 11
US-09-949-016-15419/c
; Sequence 15419, Application US/09949016

```

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL0011307
CURRENT APPLICATION NUMBER: US/09/349,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15419
LENGTH: 212449
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(212449)
OTHER INFORMATION: n = A,T,C or G

```

US-09-949-016-15419

Query Match 4.9%; Score 35.4; DB 4; Length 212449;
Best Local Similarity 57.8%; Pred. No. 13;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 123 GGGCTAAGAGTCATTCATCTTTCTTTCCCTTGGGGTCCGAGCCCGGAGCCCTCTCTT 182

DB 124084 GGTCTACATATTCATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 124025

QY 183 CCCCCGAGATTTCT 231

DB 124024 CTCCT 123976

RESULT 12

US-09-949-016-17592/c

Sequence 17592, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17592

LENGTH: 72992

TYPE: DNA

ORGANISM: Human

US-09-949-016-17592

Query Match 4.9%; Score 35.2; DB 4; Length 72992;

Best Local Similarity 60.4%; Pred. No. 7.9;

Matches 58; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 132 AGCTATCACTTTCTTTCT 191

DB 72622 AGCTCTTCT 72563

QY 192 ATTCTCTCTTCAATATCTCTCTTTATTCCTATCC 227

DB 72562 CTTTCCT 72527

RESULT 13

US-09-949-016-14133/c

Sequence 14133, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14133

LENGTH: 98962

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(98962)

OTHER INFORMATION: n = A,T,C or G

US-09-949-016-14133

Query Match 4.8%; Score 34.8; DB 4; Length 98962;

Best Local Similarity 53.7%; Pred. No. 13;

Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 84 TACTTGAAGAAAGAGCAAGCTTGTAAAGGGAGTGGGGCTTAAGAGTCATTCATT 143

DB 78525 TCCTGGGAGGACAGTGGGGTGTGAGTGGGGAGGCTTCAACCCCAACATTT 78466

QY 144 TCTTTTCCCTTCGGGTCGCGAGCCCGGAGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 203

DB 78465 TCCCAACCTTGGCCACACACCCCTGGAACCCCAAGCTCTTAAGCAATATCTTTT 78406

QY 204 CATATCTTCTTTT 217

DB 78405 TTTTCTTTTCTTTT 78392

RESULT 14

US-09-949-016-17100/c

Sequence 17100, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17100

LENGTH: 102884

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(102884)

OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17100

Query Match 4.8%; Score 34.8; DB 4; Length 102884;

Best Local Similarity 53.7%; Pred. No. 13;

Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 84 TACTTGAAGAAAGAGCAAGCTTGTAAAGGGAGTGGGGCTTAAGAGTCATTCATT 143

DB 78525 TCCTGGGAGGACAGTGGGGTGTGAGTGGGGAGGCTTCAACCCCAACATTT 78466

QY 144 TCTTTTCCCTTCGGGTCGCGAGCCCGGAGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 203

DB 78465 TCCCAACCTTGGCCACACACCCCTGGAACCCCAAGCTCTTAAGCAATATCTTTT 78406

QY 204 CATATCTTCTTTT 217

DB 78405 TTTTCTTTTCTTTT 78392

RESULT 15


```

US-09-949-016-26160/C
; Sequence 26160, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26160
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26160

```

Query Match	4.8%	Score 34.4	DB 4	Length 601
Best Local Similarity	54.8%	Pred. No. 0.82		
Matches 68, Conservative	0	Mismatches 56	Indels 0	Gaps 0

Qy	25	CACATGACATGCGTCATTAATAATACACTACGAAAAACCTTAAAGAGCAAGCATCTCT	84
	409	CAAGAGAGAGACTCTCAAAAAACAAAACAAAACAAAACAAAACAAAACGAGAGAG	350
Qy	85	ACTTGAAGAAAAAGAGACGCTTGTAAGGGGATGGGGGCTAAGAGTCATTCATTT	144
	349	ACTGGCCAGAAACAGAGTATATGAGCAAAAGAGGTGTGATGAAGAATCTTATATATGT	290
Qy	145	CTTT	148
	289	CATT	286

Search completed: August 27, 2005, 10:04:45
Job time : 182 secs

This Page Blank (uspto)


```
Db      241 GCACTATGACATAAGTGTGCTGAGACATCTCCATGCTGACTTGTGTATCTCAAGT 300
Qy      301 GGTAACGGACCGCTGGCTCGGAAACGGTTCTTCTGACAAATTCTAGAACAGGGGCTACA 360
Db      301 GGTAACGGACCGCTGGCTCGGAAACGGTTCTTCTGACAAATTCTAGAACAGGGGCTACA 360
Qy      361 GTCCTGATTAATGAAATTAATAGGCGCATTTTGTCTAGCCGCCGCCGCCGCTTTCCCA 420
Db      361 GTCCTGATTAATGAAATTAATAGGCGCATTTTGTCTAGCCGCCGCCGCCGCTTTCCCA 420
Qy      421 ATAGGAGGCGGACGTTATCGCGAGCTCTACCTTCTTCTATTTGGGTAAGCCCTTTC 480
Db      421 ATAGGAGGCGGACGTTATCGCGAGCTCTACCTTCTTCTATTTGGGTAAGCCCTTTC 480
Qy      481 TGTTCCTGCGCAAGTGTCTGAGGCTCGCCGCGAGAACATAGTGAATAGGAGTGAAC 540
Db      481 TGTTCCTGCGCAAGTGTCTGAGGCTCGCCGCGAGAACATAGTGAATAGGAGTGAAC 540
Qy      541 TTTTCATGAGAGATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 600
Db      541 TTTTCATGAGAGATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 600
Qy      601 TCATATTAAGAGAGAAATTTGCTCACTATGTGACATTTCTGGAGCTCTTAACCTT 660
Db      601 TCATATTAAGAGAGAAATTTGCTCACTATGTGACATTTCTGGAGCTCTTAACCTT 660
Qy      661 TATTGACAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATTA 720
Db      661 TATTGACAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATTA 720
Qy      721 CAT 723
Db      721 CAT 723

RESULT 2
AR492049      11427 bp      DNA      linear      PAT 15-MAY-2004
LOCUS      AR492049
DEFINITION      Sequence 24 from patent US 6716601.
ACCESSION      AR492049
VERSION      AR492049.1 GI:47260518
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 11427)
AUTHORS      Belfield,G.P. and Oakley,C.
TITLE      Compositions and methods utilizing the yeast ZEO1 promoter
JOURNAL      Patent: US 6716601-A 24 06-Apr-2004;
FEATURES
source      Location/Qualifiers
            1..11427
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 723; DB 6; Length 11427;
Best Local Similarity 100.0%; Pred. No. 6.1e-207;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTGATTAGACACGACACATCATAGACTGGCTCATTAATAATACACTACGAAAA 60
Db      15 CTTTGATTAGACACGACACATCATAGACTGGCTCATTAATAATACACTACGAAAA 74
Qy      61 ACCATTAAGAGCAAGCGATCCTACTTGGAGGAAAAAGAGACACGCTTGTAAGGGGAT 120
Db      75 ACCATTAAGAGCAAGCGATCCTACTTGGAGGAAAAAGAGACACGCTTGTAAGGGGAT 134
Qy      121 GGGGGCTAAGAAATCATCTTTCTTTCCTTCGCGGTCCGAGCCCGGGAACCCCTCCT 180
Db      135 GGGGGCTAAGAAATCATCTTTCTTTCCTTCGCGGTCCGAGCCCGGGAACCCCTCCT 194
Qy      181 CTCCTCCGACGATTTCTTCTTATATTTCTTTATTCCTTATTCGATCCGTTAAGCAAC 240

```

```
Db      195 CTCCTCCGACGATTTCTTCTTATATCTTCTTTATTTCTATCCCGTTGAAGCAAC 254
Qy      241 GCACTATGACATAATGTGTGACATCTCCATGCTGACTTGTGTATCTCAAGT 300
Db      255 GCACTATGACATAATGTGTGACATCTCCATGCTGACTTGTGTATCTCAAGT 314
Qy      301 GGTAACGGACCGCTGGCTCGGAAACGGTTCTTCTGACAAATTCTAGAACAGGGGCTACA 360
Db      315 GGTAACGGACCGCTGGCTCGGAAACGGTTCTTCTGACAAATTCTAGAACAGGGGCTACA 374
Qy      361 GTCCTGATTAATGAAATTAATAGGCGCATTTTGTCTAGCCGCCGCCGCCGCTTTCCCA 420
Db      375 GTCCTGATTAATGAAATTAATAGGCGCATTTTGTCTAGCCGCCGCCGCCGCTTTCCCA 434
Qy      421 ATAGGAGGCGGACGTTATCGCGAGCTCTACCTTCTTCTATTTGGGTAAGCCCTTTC 480
Db      435 ATAGGAGGCGGACGTTATCGCGAGCTCTACCTTCTTCTATTTGGGTAAGCCCTTTC 494
Qy      481 TGTTCCTGCGCAAGTGTCTGAGGCTCGCCGCGAGAACATAGTGAATAGGAGTGAAC 540
Db      495 TGTTCCTGCGCAAGTGTCTGAGGCTCGCCGCGAGAACATAGTGAATAGGAGTGAAC 554
Qy      541 TTTTCATGAGAGATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 600
Db      555 TTTTCATGAGAGATTAGCAAGCGGAAAAAACTATGCTAGCTGGAGTTGTTTTCAA 614
Qy      601 TCATATTAAGAGAGAAATTTGCTCACTATGTGACATTTCTGGAGCTCTTAACCTT 660
Db      615 TCATATTAAGAGAGAAATTTGCTCACTATGTGACATTTCTGGAGCTCTTAACCTT 674
Qy      661 TATTGACAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATTA 720
Db      675 TATTGACAGAGACTATCAATCATACAGATATTGTCAAAAAAAAGACTAATATTA 734
Qy      721 CAT 723
Db      735 CAT 737

RESULT 3
AR492045      13073 bp      DNA      linear      PAT 15-MAY-2004
LOCUS      AR492045
DEFINITION      Sequence 20 from patent US 6716601.
ACCESSION      AR492045
VERSION      AR492045.1 GI:47260514
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 13073)
AUTHORS      Belfield,G.P. and Oakley,C.
TITLE      Compositions and methods utilizing the yeast ZEO1 promoter
JOURNAL      Patent: US 6716601-A 20 06-Apr-2004;
FEATURES
source      Location/Qualifiers
            1..13073
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 723; DB 6; Length 13073;
Best Local Similarity 100.0%; Pred. No. 6.3e-207;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CTTTGATTAGACACGACACATCATAGACTGGCTCATTAATAATACACTACGAAAA 60
Db      16 CTTTGATTAGACACGACACATCATAGACTGGCTCATTAATAATACACTACGAAAA 75
Qy      61 ACCATTAAGAGCAAGCGATCCTACTTGGAGGAAAAAGAGACACGCTTGTAAGGGGAT 120
Db      76 ACCATTAAGAGCAAGCGATCCTACTTGGAGGAAAAAGAGACACGCTTGTAAGGGGAT 135
Qy      121 GGGGGCTAAGAAATCATCTTTCTTTCCTTCGCGGTCCGAGCCCGGGAACCCCTCCT 180

```

Db 136 GGGGGCTAGAGATCATTCATCTTTTCCCTTCGGCGTCCGAGCCCGGACCCCTCCT 195
QY 181 CTCGCCGACAGATTTCTTCTTATCATATCTTCTTTTATTCATTCCTGTAAGAACCC 240
Db 196 CTCGCCGACAGATTTCTTCTTATCATATCTTCTTTTATTCATTCCTGTAAGAACCC 255
QY 241 GCACATAGCTAAATGGTGTCTGACATCTCCATGGCTGTGATCTGTATCTCAAGT 300
Db 256 GCACATAGCTAAATGGTGTCTGACATCTCCATGGCTGTGATCTGTATCTCAAGT 315
QY 301 GGTAAAGGACCGTGTCTGGAACCGTTCCTTCGTGACAAATTCGTAAAGGGCTACA 360
Db 316 GGTAAAGGACCGTGTCTGGAACCGTTCCTTCGTGACAAATTCGTAAAGGGCTACA 375
QY 361 GTCGTATATAGAAATTAATAGGCAATTTTGTAGAGCCGCGCGCCGCTTTCCCA 420
Db 376 GTCGTATATAGAAATTAATAGGCAATTTTGTAGAGCCGCGCGCCGCTTTCCCA 435
QY 421 ATAGGAGGCGGAGTTTATCGGCGAGCTTACTTCTTCTATTTGGTAAAGCCCTTTC 480
Db 436 ATAGGAGGCGGAGTTTATCGGCGAGCTTACTTCTTCTATTTGGTAAAGCCCTTTC 495
QY 481 TGTTCCTGCGCAATGTGTCTGACAGCTGCGCGGAGAAATAGTAAAGGATGTAAAC 540
Db 496 TGTTCCTGCGCAATGTGTCTGACAGCTGCGCGGAGAAATAGTAAAGGATGTAAAC 555
QY 541 TTTGCATGAGAGAAATTAAGCAAGCGGAAATACTATGGCTAGGAGTGTTTTCAA 600
Db 556 TTTGCATGAGAGAAATTAAGCAAGCGGAAATACTATGGCTAGGAGTGTTTTCAA 615
QY 601 TCATATTAAGAGGAGAAATTTGCTCCTCATATGTGACAGTTTCGGAGCTTTAACTTT 660
Db 616 TCATATTAAGAGGAGAAATTTGCTCCTCATATGTGACAGTTTCGGAGCTTTAACTTT 675
QY 661 TATTCGAGAGGACTATCAATCATACAGATATGTCTCAAAAAAAGACTAATATATA 720
Db 676 TATTCGAGAGGACTATCAATCATACAGATATGTCTCAAAAAAAGACTAATATATA 735
QY 721 CAT 723
Db 736 CAT 738

RESULT 4
AR492055 850 bp DNA linear PAT 15-MAY-2004
DEFINITION Sequence 30 from patent US 6716601.
ACCESSION AR492055
VERSION AR492055.1 GI:47260524
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 850)
AUTHORS Belfield,G.P. and Oakley,C.
TITLES Compositions and methods utilizing the yeast ZEO1 promoter
JOURNAL Patent: US 6716601-A 30 06-APR-2004;
FEATURES
location/Qualifiers
1..850
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 99.6%; Score 720.4; DB 6; Length 850;
Best Local Similarity 99.9%; Pred. No. 2,4e-206;
Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTTTCGATTGACGACACATCATCATGAGTGGTCAATAAATACACTAGGAAAA 60
Db 78 CTTTCGATTGACGACACATCATCATGAGTGGTCAATAAATACACTAGGAAAA 137
QY 61 ACCATAAAGCAAGGATACCTACTTGGAGGAAAGAGGACCGCTTGTAAAGGGGAT 120

Db 138 ACCATAAAGCAAGGATACCTACTTGGAGGAAAGAGGACCGCTTGTAAAGGGGAT 197
QY 121 GGGGGCTAAGAGTCAATTCATTTCTTTCCCTTCGCGGTCCGAGCCCGGACCCCTCCT 180
Db 198 GGGGGCTAAGAGTCAATTCATTTCTTTCCCTTCGCGGTCCGAGCCCGGACCCCTCCT 257
QY 181 CTCGCCGACAGATTTCTTCTTATCATATCTTCTTTTATTCATTCCTGTAAGAACCC 240
Db 256 CTCGCCGACAGATTTCTTCTTATCATATCTTCTTTTATTCATTCCTGTAAGAACCC 317
QY 241 GCACATAGCTAAATGGTGTCTGACATCTCCATGGCTGTGATCTGTATCTCAAGT 300
Db 316 GGTAAAGGACCGTGTCTGGAACCGTTCCTTCGTGACAAATTCGTAAAGGGCTACA 377
QY 301 GGTAAAGGACCGTGTCTGGAACCGTTCCTTCGTGACAAATTCGTAAAGGGCTACA 360
Db 376 GGTAAAGGACCGTGTCTGGAACCGTTCCTTCGTGACAAATTCGTAAAGGGCTACA 437
QY 421 ATAGGAGGCGGAGTTTATCGGCGAGCTTACTTCTTCTATTTGGTAAAGCCCTTTC 480
Db 436 ATAGGAGGCGGAGTTTATCGGCGAGCTTACTTCTTCTATTTGGTAAAGCCCTTTC 497
QY 481 TGTTCCTGCGCAATGTGTCTGACAGCTGCGCGGAGAAATAGTAAAGGATGTAAAC 540
Db 496 TGTTCCTGCGCAATGTGTCTGACAGCTGCGCGGAGAAATAGTAAAGGATGTAAAC 557
QY 541 TTTGCATGAGAGAAATTAAGCAAGCGGAAATACTATGGCTAGGAGTGTTTTCAA 600
Db 556 TTTGCATGAGAGAAATTAAGCAAGCGGAAATACTATGGCTAGGAGTGTTTTCAA 617
QY 601 TCATATTAAGAGGAGAAATTTGCTCCTCATATGTGACAGTTTCGGAGCTTTAACTTT 660
Db 618 TTTGCATGAGAGAAATTAAGCAAGCGGAAATACTATGGCTAGGAGTGTTTTCAA 677
QY 661 TATTCGAGAGGACTATCAATCATACAGATATGTCTCAAAAAAAGACTAATATATA 720
Db 678 TATTCGAGAGGACTATCAATCATACAGATATGTCTCAAAAAAAGACTAATATATA 737
QY 721 CA 722
Db 738 TATTCGAGAGGACTATCAATCATACAGATATGTCTCAAAAAAAGACTAATATATA 797
QY 798 AA 799

RESULT 5
SC9920 23498 bp DNA linear PLN 11-AUG-1997
LOCUS S.cerevisiae chromosome XIII cosmid 9920.
DEFINITION Z48639 Z71257
ACCESSION Z48639.1 GI:732924
VERSION Z48639.1 GI:732924
KEYWORDS COX7; cytochrome oxidase; delta element; glutamate decarboxylase;
PEPtil; transfer RNA-Ala.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
REFERENCE 1 (bases 1 to 23498)
AUTHORS Hunt,S. and Bowman,S.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 23498)
AUTHORS Bartell,B., Rajandream,M.A. and Walsh,S.V.
TITLES Direct Submision
JOURNAL Submitted (10-MAR-1995) Saccharomyces cerevisiae chromosome XIII
sequencing project, Sanger Centre, Hinxton Hall, Hinxton, Cambridge
CB10 1RQ E-mail: bartell@sanger.ac.uk

NOTES:
All CDS over 100 codons have been analysed. CDS that are completely
overlapped and those that are overlapped by more than 50%
of their length by a larger CDS have been omitted from this
analysis.

Details of the omitted CDS are available on request. The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons and the calculated codon adaptation index (CAI) is given for each CDS.
Cosmid 9920 is overlapped at the start of this sequence by cosmid 9408

and at the end of this sequence by cosmid 8156.

Location/Qualifiers

1. 23498

FEATURES
source

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/strain="AB972"

/db_xref="taxon:4932"

/chromosome="XIII"

/map="13R"

/clone="cosmid 9920"

complement(41..2870)

/note="YM9920.01c, unknown, partial, len: 956, CAI: 0.14;

P00061 Short-chain alcohol dehydrogenase family

signature

/codon_start=1

/product="unknown"

/protein_id="CA88574.1"

/db_xref="GI:732925"

/translation="MSFGINTFOQYNTDGLGHNGVRIISLNYFDGLPDPSSLNLSYS

NELKIFKILKREDTTEKALMDLSNLISDPNOEYFPNDIFLQMSQIYAKLIISD

YVIRLQSHOITIMLVKSLRKIKSKLPFLILGTCGLDYSVKPSKINLTCPN

KPKAKINLMAVFOQLNLVKEIYVNEEDTISDRYSKESSEPRHRYVASAVL

LKLFVHNQDVERNSSSLKVLSDESIYKLNKLNQGNINAYETVRLIDVLYRGY

MSHKINIKLAVKLIKLSLTHITSKINLVCPLPSILNLATLDDYEDGITWSYDKS

SEKRYLFLVSRTSPSPGFNAVFALYSTKHSFLDYLLMLPFMQSKVQLNKG

FSARNSAEVLNEFWTNPLFAEDSSEERVKWVESIPNSLSCGSLSYTLNQTLS

GVPPDPKREIIEYFTSDSDIRKIVSEKXLPALLIYSPNNESAISRLPFPVOLI

ENDSVNPKYGVYDALNYFLDSMIPLANGIKRTIKRTINPTLVNVESTIQNRAGIMQ

YNSKRFKONTDAITSLEDFTVALSFINPKTITLATNMLNDIYQQLMKSDELE

LVIIEPMKNYKFDSDGEIPIKGNKFLNQTITTLVRSAVANQVEQFCVLSKDETF

FSTLLINDFLCALYEVSIEDNTEKLFKLSQLAKGNSIANKLQVILIQHAYFSP

GAKERYTHAVELINGCNDTSQIPFPANAIEVARYMPAIDYRSSIVSSISTNTILL

TDDPNTLNKMKILRYALFLDALDALPERVNNHVARITVVSSELVTNCLSEPN

DIYDPGHTFPFKGKNLNFSDIVGVNVIOPANGDAMLFDIAESNSVFFYYSKVL

KVLNSIDTVSSSTTLNGLASVSPVTKVTRQKSTDKOYL"

complement(2406..2438)

/note="P00061 Short-chain alcohol dehydrogenase family

signature"

/product="unknown"

/protein_id="CA88574.1"

/db_xref="GI:732925"

/translation="MSFGINTFOQYNTDGLGHNGVRIISLNYFDGLPDPSSLNLSYS

NELKIFKILKREDTTEKALMDLSNLISDPNOEYFPNDIFLQMSQIYAKLIISD

YVIRLQSHOITIMLVKSLRKIKSKLPFLILGTCGLDYSVKPSKINLTCPN

KPKAKINLMAVFOQLNLVKEIYVNEEDTISDRYSKESSEPRHRYVASAVL

LKLFVHNQDVERNSSSLKVLSDESIYKLNKLNQGNINAYETVRLIDVLYRGY

MSHKINIKLAVKLIKLSLTHITSKINLVCPLPSILNLATLDDYEDGITWSYDKS

SEKRYLFLVSRTSPSPGFNAVFALYSTKHSFLDYLLMLPFMQSKVQLNKG

FSARNSAEVLNEFWTNPLFAEDSSEERVKWVESIPNSLSCGSLSYTLNQTLS

GVPPDPKREIIEYFTSDSDIRKIVSEKXLPALLIYSPNNESAISRLPFPVOLI

ENDSVNPKYGVYDALNYFLDSMIPLANGIKRTIKRTINPTLVNVESTIQNRAGIMQ

YNSKRFKONTDAITSLEDFTVALSFINPKTITLATNMLNDIYQQLMKSDELE

LVIIEPMKNYKFDSDGEIPIKGNKFLNQTITTLVRSAVANQVEQFCVLSKDETF

FSTLLINDFLCALYEVSIEDNTEKLFKLSQLAKGNSIANKLQVILIQHAYFSP

GAKERYTHAVELINGCNDTSQIPFPANAIEVARYMPAIDYRSSIVSSISTNTILL

TDDPNTLNKMKILRYALFLDALDALPERVNNHVARITVVSSELVTNCLSEPN

DIYDPGHTFPFKGKNLNFSDIVGVNVIOPANGDAMLFDIAESNSVFFYYSKVL

KVLNSIDTVSSSTTLNGLASVSPVTKVTRQKSTDKOYL"

complement(2406..2438)

/note="P00061 Short-chain alcohol dehydrogenase family

signature"

/product="unknown"

/protein_id="CA88574.1"

/db_xref="GI:732926"

/translation="MGEIAPNKPDPBGTINMCOQLRYIYMMSIGLRLLPLNKAII

IFLNSLALNLCALITYE"

complement(3014101..4196,4434..4505)

/note="YM9920.03c, unknown, len: 55, CAI: 0.13, possible

small spliced gene"

/codon_start=1

/product="unknown"

/protein_id="CA88576.1"

/db_xref="GI:732927"

/translation="MELMASCFRMPRPRYRANPTLGLWLTGLRLPLNKAIIIPNSL

ALAHCAIYTE"

complement(4197..4199)

/note="possible splice acceptor sequence, caag"

misc_feature
misc_feature
misc_feature
misc_feature

CDS

complement(4222..4228)
/note="possible splice branch sequence, taccac"
complement(4428..4433)
/note="possible donor sequence, gtagt"
complement(4685..4690)
/note="possible donor sequence, gtagt"
complement(4689..4694)
/note="possible donor sequence, gtagt"
5632..7389
/note="YM9920.04, unknown, len: 585, CAI: 0.18, putative
glutamate decarboxylase gene"

/codon_start=1

/product="unknown"

/protein_id="CA88577.1"

/db_xref="GI:732928"

/db_xref="GOA:004792"

/db_xref="SGD:50004862"

/db_xref="UniProt/Swiss-Prot:Q04792"

/translation="MLHRSKQKNFENIACKVYHDLAQLDLSNDVOKSAVQSGHOG

SNMMDTSSQGMANKYSVPKGLPADLSYQLIHNELTLDGPHNLASFYVFTTDQA

RKLIDENLTKNLADNDEYPLIETLORCISMLAQIMHNPDEEPIGCATGSSSEAIML

GLAMKKRHERMKAGADASKPNIINSSACQVALKEPTRYFEVSCRIPVSHSHM

LDPSLMDYVDENTIGCFVILGTYTGHLENVEKADYLSQIEAKHPMSNTDPIHA

DGASGGFTIIPRPFKEHKAYGMERMGFNHRRVVSMTSGHKFGITTPGLMVLWDE

SLADELAFKIKYIGVETFGILNPSRPGVYVNIQYRPVSLGSHGYTFOQNSLFA

RAFSFELNSSLKPCCEFISSIHESIINDSPASVKDYMHPQAYKGPLVAFKLS

KKFFHEYEVEVQALLSLRGRGWIPVYPLPKATDSDSEKERVRRVRSBMKDLAQ

LLIVDIESITLTKLHSYEKCHHIELASEQPERKSSPIYEMLLALASPQDDIIPDE

IEKNKLEWTRHRYRC"

7746..8846

/note="YM9920.05, unknown, len: 366, CAI: 0.11, similar to

YK56_YEST P36156, YKR076w, hypothetical protein in S152,

62.6% identity in 369 aa overlap"

/codon_start=1

/product="unknown"

/protein_id="CA88578.1"

/db_xref="GI:732929"

/db_xref="GI:50004863"

/db_xref="UniProt/Swiss-Prot:Q04806"

/translation="MSEKASANNKAEFRQSSPPREIISADHPITYKPAKGRWYVAL

PCPMARLITRALKGLAPITIGCSVAHWHLDKMGKFLSEDDGKTNEBHWFIDGIS

SVNNTSPVANINIPNNARLLVDGDEPHGVKRLSDPEFKTDPKGRFTVPVLMG

ETCTIVNESSDITIGMNSAAPDERVEGRYQVALVPSLEAQTTERPSWYDKING

VYKGFACCAEVREVSFLFYLDKLENLDDKTYTDEALYKNNKOKITDRIPAIG

DTLTEADRYLPTIVRPDVVYHOFKMLAIIIRDYSRHTWLNKNIYRHEAFORTTD

FTHIKLYTRSQPRVNPIGITPLGPKDIRP"

complement(10146..10550)

/note="YM9920.06c, unknown, len: 134, CAI: 0.13"

/codon_start=1

/product="unknown"

/protein_id="CA88579.1"

/db_xref="GI:732930"

/db_xref="SGD:50004865"

/db_xref="UniProt/Swiss-Prot:Q04814"

/translation="MFGKVFYSYIRTRIGRFLPSITVYVSSSISPDKEACFPKKY

HELNNSQKEFTQRPVKNYRHQYSSKTNVSLKGISIMDEHNDSPSVFGIFYNDIWK

SPKNEQLGTNDNKKSGSRFSHPFKOLLQK"

complement(10777..12021)

/note="YM9920.07c, unknown, len: 414, CAI: 0.13,

hydrophobic domains"

/codon_start=1

/product="unknown"

/protein_id="CA88580.1"

/db_xref="GI:732931"

/db_xref="GOA:004835"

/db_xref="SGD:50004866"

/db_xref="UniProt/Swiss-Prot:Q04835"

/translation="MNSVPVKMRKENTHLLVSKENDTSLQLPSTTRISLSPKESNS

NEDFVNDNETTLQRIKSDYLPNIGVLLSVFSSKAVMSKTVLENDPDDIANDR

OIKPQLIYLVAVITIGTLIYNIKSTLSDVDFGKEXKMKVLRCTGDFGANDR

YSLMPLTDAVLIITFLAPSLTIFLSMILIREFTKVEALGSLISLGVALLVPSF

LGRTPELTDSSQVIVESDPSKRLIATILYVIGMVGMSCVYIIIRYIGRAIAMSVS

YFSLITAVSFTGINTIPSMKFOIPHSKKQWILFENLVSGFTIQLLTMTGIQRBAC

misc_feature

/note="possible splice acceptor sequence, caag"

CDS
 RSLMTYQLLYAVFWDVALYKHPNIMSWIGMIIISATLWVIRIRANNETXAKDL
 TPIIDDENSIPLIEFDLSDSK"
 complement (12446..12754)
 /note="Ym9920.08c, unknown, len: 103, CAI: 0.07,
 questionable orf"
 /codon_start=1
 /product="unknown"
 /protein_id="CAA8581.1"
 /db_xref="GI:732932"
 /db_xref="GCA:004838"
 /db_xref="SGD:S0004867"
 /db_xref="UniProt/Swiss-Prot:Q04838"
 /translation="MVPLILILFLFSKFSFTLPVNVHVLTKYTAIVNTKMQTPSII
 DVLYTHVHFYMTIILIVRKQMSIHAFISLCLPSHVLDFSIVRDILSMYFLETAV
 "

CDS
 12832..13398
 /note="Ym9920.09, unknown, len: 188, CAI: 0.15,
 colied-col1 domain"
 /codon_start=1
 /product="unknown"
 /protein_id="CAA8582.1"
 /db_xref="GI:732933"
 /db_xref="GCA:004839"
 /db_xref="SGD:S0004868"
 /db_xref="UniProt/Swiss-Prot:Q04839"
 /translation="MPLSIVADAPDEEPKPKPSHKRSNNKNNKNNNSWNSNSSN
 NKKDSVNVKNNKNNHESKTKIKETLPREKKPHSQGKISPSSESLAIPFSQKA
 TEISPPVPSPKMTTKTKOSKQDPSKMLKKKIEQREIIOKTHHKKQOOVLMDP
 LNDESSNMVDDDEELIQRLKTSIKT"
 complement (13776..13958)

gene

Query Match 99.6%; Score 720.4; DB 8; Length 23498;
 Best Local Similarity 99.9%; Pred. No. 4.2e-206;
 Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTTTCGATTAGCAGCACAATCACTAGATGCGTCAATAAAATACACTACGAAAA 60
 DB 8860 CTTTCGATTAGCAGCACAATCACTAGATGCGTCAATAAAATACACTACGAAAA 8919
 QY 61 ACCATAAGAGCAAGGATACCTACTTGGAGAGAAAGAGACCGCTTGTAAAGGGAGAT 120
 DB 8920 ACCATAAGAGCAAGGATACCTACTTGGAGAGAAAGAGACCGCTTGTAAAGGGAGAT 8979
 QY 121 GGGGGCTAAGAGTCACTTCTTTTCCCTCGGGGCGGAGCCGGGACCCCTCCT 180
 DB 8980 GGGGGCTAAGAGTCACTTCTTTTCCCTCGGGGCGGAGCCGGGACCCCTCCT 9039
 QY 181 CTCGCCGACGATTTCTTCTTATATCTTCTTTATTCCTATCCGTTGAAGCAAC 240
 DB 9040 CTCGCCGACGATTTCTTCTTATATCTTCTTTATTCCTATCCGTTGAAGCAAC 9099
 QY 241 GCACTATAGTAATGGTGTGATCTTCATAGCTGTGATCTTGTGTATCTCAAGT 300
 DB 9100 GCACTATAGTAATGGTGTGATCTTCATAGCTGTGATCTTGTGTATCTCAAGT 9159
 QY 301 GGTAAAGGACCGTGTGCGGAAACGTTCTTCGGAATCTCTGAACAAGGGGCTACA 360
 DB 9160 GGTAAAGGACCGTGTGCGGAAACGTTCTTCGGAATCTCTGAACAAGGGGCTACA 9219
 QY 361 GTCTCGATATAGTAATAGGCAATTTTGTAGAGCGCGCGGCGCGCTTTCCCA 420
 DB 9220 GTCTCGATATAGTAATAGGCAATTTTGTAGAGCGCGCGCGCGCTTTCCCA 9279
 QY 421 ATAGGAGGCGAGTTATCGGCGAGCTCTTCTTCTATTTGGGTAAGCCCTTTTC 480
 DB 9280 ATAGGAGGCGAGTTATCGGCGAGCTCTTCTTCTATTTGGGTAAGCCCTTTTC 9339
 QY 481 TGTTCGCGCAAGTGTGTGAGGCTGCGCGGAGAAATAGTATAAGGATGTAAC 540
 DB 9340 TGTTCGCGCAAGTGTGTGAGGCTGCGCGGAGAAATAGTATAAGGATGTAAC 9399
 QY 541 TTTGATGAGATTAAGCAAGGAAAAAATAGCTAGCTGAGGAGTTGTTTTTCAA 600

DB 9400 TTTTCATGAGAGATTAGCAACGGAAAAAACTATGGCTAGCTGGAGCTTTTTCAA 9459
 QY 601 TCATATTAAGAGAGAAATTTGTGCTCATATGTGACAGTTTCTGGAGCTTAACTTT 660
 DB 9460 TCATATTAAGAGAGAAATTTGTGCTCATATGTGACAGTTTCTGGAGCTTAACTTT 9519
 QY 661 TATTCAGAGAGCTATCAATCATACAGATATTGTCAAAAAAAGACTPAATATA 720
 DB 9520 TATTCAGAGAGCTATCAATCATACAGATATTGTCAAAAAAAGACTPAATATA 9579
 QY 721 CA 722
 DB 9580 AA 9581

RESULT 6
 AX536716 680 bp DNA linear PAT 22-NOV-2002
 LOCUS
 DEFINITION Sequence 317 from Patent WO2064766.
 AX536716
 ACCSSION
 VERSION AX536716.1 GI:25263152
 KEYWORDS
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae
 Bacteria; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetaceae; Saccharomycetaceae; Saccharomyces.

REFERENCE
 1 Contreras, R.H., Eberhardt, I., Luyten, W.H. and Reekmans, R.J.
 Bax-responsive genes for drug target identification in yeast and
 fungi
 Patent: WO 02064766-A 317 22-AUG-2002;
 JANSSEN PHARMACEUTICA N.V. (BE)
 location/Qualifiers
 1..680
 /organism="Saccharomyces cerevisiae"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4932"

ORIGIN

Query Match 68.8%; Score 497.4; DB 6; Length 680;
 Best Local Similarity 99.8%; Pred. No. 6.7e-139;
 Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 224 ATCCGTTAAGCAACCGCATATGATGATGCTGATCATCTGCTGAT 283
 DB 1 ATCCGTTAAGCAACCGCATATGATGATGCTGATCATCTGCTGAT 60
 QY 284 TGTGTATCTCAGATGTGTAACGCAACCGTGGCTCGAAACGTTCTTGTGCAAT 343
 DB 61 TGTGTATCTCAGATGTGTAACGCAACCGTGGCTCGAAACGTTCTTGTGCAAT 120
 QY 344 CTAGAACAGGGCTACAGTCTCGATATATGATATTAAGCAATTTTCTAGCGCGCC 403
 DB 121 CTAGAACAGGGCTACAGTCTCGATATATGATATTAAGCAATTTTCTAGCGCGCC 180
 QY 404 GGGGGCCCGTTTCCCAATAGGAGGCGAGTTATCGCGGAGCTCTACTTCTTCTAT 463
 DB 181 GGGGGCCCGTTTCCCAATAGGAGGCGAGTTATCGCGGAGCTCTACTTCTTCTAT 240
 QY 464 TTGGGTAAGCCCTTCTGTGTTTCGCGCAGTGTGCTGCAAGCTGCGCGGAGCAATA 523
 DB 241 TTGGGTAAGCCCTTCTGTGTTTCGCGCAGTGTGCTGCAAGCTGCGCGGAGCAATA 300
 QY 524 GTGATAGGAGATGTAATTTGATAGAGATTTAGCAAGCGGAAAAAATATGCTTAC 583
 DB 301 GTGATAGGAGATGTAATTTGATAGAGATTTAGCAAGCGGAAAAAATATGCTTAC 360
 QY 584 TGGAGTGTGTTTCAATCATATAAGGAGAAATTTGTGCTCATATGTGACAGTTTC 643
 DB 361 TGGAGTGTGTTTCAATCATATAAGGAGAAATTTGTGCTCATATGTGACAGTTTC 420
 QY 644 TGGAGTGTGTTTCAATCATATAAGGAGAAATTTGTGCTCATATGTGACAGTTTC 703

repeat_region	QADQLPMADKTSARHVVDQSNKRVN"	4159. .4258
repeat_region	/rpt_type=candem	4159. .4258
repeat_region	4159. .4215	
repeat_region	/rpt_type=candem	4718. .4749
repeat_region	/rpt_type=candem	4922. .4978
repeat_region	/rpt_type=candem	5175. .5199
repeat_region	/rpt_type=candem	5322. .5352
repeat_region	/rpt_type=candem	5350. .5397
repeat_region	/rpt_type=candem	5405. .5438
repeat_region	/rpt_type=candem	5531. .5557
gene	/rpt_type=candem	<5712. .>6914
mrna	/locus_tag="Pp11_0072"	join(<5712. .5802,5896. .6191,6311. .6454,6636. .>6914)
cds	/locus_tag="Pp11_0072"	join(5712. .5802,5896. .6191,6311. .6454,6636. .6914)
	/locus_tag="Pp11_0072"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="AAN35661.1"	
	/db_xref="GI:23495999"	
	/translation="MKAKIKSCVTCCLFFPIYQIWINCLPFCGVNINKNLGYPSPLIK	
	KNVQEKPKRQLKVQNEKENIPTQDGNNTI KEDTKRKQTNTINDINKLEKKIEI	
	NEIKKSNINILERSFSGFHKEDITTPBELLKELIQYRKIFQRHDKCGSSBIOIIL	
	LTPIKPTLETHMKKKDKPACLRGLFCVSKRRRLVYGRKDRREMPKQTSYNIKK	
	PLRPTREYVYKDLKYIHFNNTGRFKNNAKKKKKKXKLKKKKYQDIXLFGN"	
repeat_region	6250. .6288	
repeat_region	/rpt_type=candem	6482. .6530
repeat_region	/rpt_type=candem	6482. .6530
repeat_region	/rpt_type=candem	6484. .6533
repeat_region	/rpt_type=candem	6913. .6938
repeat_region	/rpt_type=candem	7132. .7156
repeat_region	/rpt_type=candem	7287. .7327
repeat_region	/rpt_type=candem	7691. .7726
gene	/rpt_type=candem	<7800. .>10244
mrna	/locus_tag="Pp11_0073"	join(<7800. .9712,9978. .10049,10155. .10162,10204. .>10244)
cds	/locus_tag="Pp11_0073"	join(7800. .9712,9978. .10049,10155. .10162,10204. .10244)
	/locus_tag="Pp11_0073"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="AAN35662.1"	
	/db_xref="GI:23496000"	
	/translation="WFCSPFLFFPFPVQNKELIALSCVHLKKKHIVPRKSGHPI	
	YSVDNNVNLKNKNEYSKLKRIKIYEQVHNKKCKIYPLPIFLKNNNDIYEDH	
	KDKIKDNNNNKKKKDKDIDNNNDKKKINDTPAONITNDKXKNSNFEDVYEDH	
	LKNNILVESSYINKSDVLIYPLNHYDILNKLLEIIPNNIINSLDYIFLNKKDYNK	
	SDIINNNNNNNNNNKYNICPEYNNIYNIIFPIYVNDQDNPINISGYSPEFTNK	
	DNIIHKRPOLVYIQKLENDHYNLKQDQEHKKYINIPLEPKNIIDINKIDQIEE	
	NKMLYPRKNIKLVKEINVTNTNTNTNKQEDQNLIKYVHKNDTPFKYKPTLKN	
	LITNEVYIYVFLSSYKDLIKETIYIAKOTTIPNNVIKHIONINIKKQHNEILLN	
	NFKQKINDKINSNVNQVEIKKKQYSYLHIIISDLLEKYSNEILKKNLYVQFKEN	
	IOKIKINDSILQFKKQIINHIDQLFDYNSKIFILPSNNKYOQKQSYIYYKKIKL	
	LQILNQNPAGSEINVIYISGLYKNGYSFNFESKCYKTLINVIYITIFILNKNKNINE	
	NYLSPNMGESYVDDISLSPKKIDIMTTSEMDYQKIYIQISLVVPRPKRGS"	
repeat_region	8865. .8890	
repeat_region	/rpt_type=candem	9936. .9975

Query Match	Best Local Match	Similarity	Score	Pos. 1	DB 3	Length	257757		
Matches	120	Conservative	0	Mismatches	102	Indels	2	Gaps	1
repeat_region	repeat_region	11438	.11479	/rpt_type=tandem					
repeat_region	repeat_region	12185	.12400	/rpt_type=tandem					
repeat_region	repeat_region	12784	.12816	/rpt_type=tandem					
repeat_region	repeat_region	13620	.13652	/rpt_type=tandem					
repeat_region	repeat_region	13757	.13788	/rpt_type=tandem					
repeat_region	repeat_region	14147	.14179	/rpt_type=tandem					
repeat_region	repeat_region	14185	.14200	/rpt_type=tandem					
repeat_region	repeat_region	14784	.14816	/rpt_type=tandem					
repeat_region	repeat_region	15620	.15652	/rpt_type=tandem					
repeat_region	repeat_region	15757	.15788	/rpt_type=tandem					
repeat_region	repeat_region	16147	.16179	/rpt_type=tandem					
repeat_region	repeat_region	16185	.16200	/rpt_type=tandem					
repeat_region	repeat_region	16784	.16816	/rpt_type=tandem					
repeat_region	repeat_region	17620	.17652	/rpt_type=tandem					
repeat_region	repeat_region	17757	.17788	/rpt_type=tandem					
repeat_region	repeat_region	18147	.18179	/rpt_type=tandem					
repeat_region	repeat_region	18185	.18200	/rpt_type=tandem					
repeat_region	repeat_region	18784	.18816	/rpt_type=tandem					
repeat_region	repeat_region	19620	.19652	/rpt_type=tandem					
repeat_region	repeat_region	19757	.19788	/rpt_type=tandem					
repeat_region	repeat_region	20147	.20179	/rpt_type=tandem					
repeat_region	repeat_region	20185	.20200	/rpt_type=tandem					
repeat_region	repeat_region	20784	.20816	/rpt_type=tandem					
repeat_region	repeat_region	21620	.21652	/rpt_type=tandem					
repeat_region	repeat_region	21757	.21788	/rpt_type=tandem					
repeat_region	repeat_region	22147	.22179	/rpt_type=tandem					
repeat_region	repeat_region	22185	.22200	/rpt_type=tandem					
repeat_region	repeat_region	22784	.22816	/rpt_type=tandem					
repeat_region	repeat_region	23620	.23652	/rpt_type=tandem					
repeat_region	repeat_region	23757	.23788	/rpt_type=tandem					
repeat_region	repeat_region	24147	.24179	/rpt_type=tandem					
repeat_region	repeat_region	24185	.24200	/rpt_type=tandem					
repeat_region	repeat_region	24784	.24816	/rpt_type=tandem					
repeat_region	repeat_region	25620	.25652	/rpt_type=tandem					
repeat_region	repeat_region	25757	.25788	/rpt_type=tandem					
repeat_region	repeat_region	26147	.26179	/rpt_type=tandem					
repeat_region	repeat_region	26185	.26200	/rpt_type=tandem					
repeat_region	repeat_region	26784	.26816	/rpt_type=tandem					
repeat_region	repeat_region	27620	.27652	/rpt_type=tandem					
repeat_region	repeat_region	27757	.27788	/rpt_type=tandem					
repeat_region	repeat_region	28147	.28179	/rpt_type=tandem					
repeat_region	repeat_region	28185	.28200	/rpt_type=tandem					
repeat_region	repeat_region	28784	.28816	/rpt_type=tandem					
repeat_region	repeat_region	29620	.29652	/rpt_type=tandem					
repeat_region	repeat_region	29757	.29788	/rpt_type=tandem					
repeat_region	repeat_region	30147	.30179	/rpt_type=tandem					
repeat_region	repeat_region	30185	.30200	/rpt_type=tandem					
repeat_region	repeat_region	30784	.30816	/rpt_type=tandem					
repeat_region	repeat_region	31620	.31652	/rpt_type=tandem					
repeat_region	repeat_region	31757	.31788	/rpt_type=tandem					
repeat_region	repeat_region	32147	.32179	/rpt_type=tandem					
repeat_region	repeat_region	32185	.32200	/rpt_type=tandem					
repeat_region	repeat_region	32784	.32816	/rpt_type=tandem					
repeat_region	repeat_region	33620	.33652	/rpt_type=tandem					
repeat_region	repeat_region	33757	.33788	/rpt_type=tandem					
repeat_region	repeat_region	34147	.34179						

Tateno, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-291G2, 1984 bp overlap; the clone sequenced to the right is RP11-158I8. Actual start of this clone is at base position 105122 of RP11-291G2; actual end is at base position 171265 of RP11-107.

Data from AC013461, AC073465, AC011667 and AC012052 was used to finish this clone, AC015764.

Polymorphisms have been identified between AC013461, AC015764 and AC012052.

FEATURES

source

The sequence of AC015764 has been incorporated into AC092573.

Location/Qualifiers

```
1..171265
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-107"
/clone_lib="RPCI-11"
570..784
/rpc_family="MIR"
1544..1678
/rpc_family="Alu"
1678..1738
/rpc_family="(TAGA)n"
1813..1835
/rpc_family="AT_rich"
2054..2351
/rpc_family="Alu"
2173..2188
/note="similar to Homo sapiens EST AL524174
(NID:g12787667)"
2352..2384
/rpc_family="L1"
2385..3107
/rpc_family="L1"
3117..3138
/rpc_family="AT_rich"
3120..3587
/rpc_family="MER1_type"
3588..3651
/rpc_family="MER1_type"
3877..3996
/note="match to EST BB867724 (NID:g10316500)"
3947..3996
/note="match to EST AL537636 (NID:g12801129)"
3956..3996
/note="match to EST AW814670 (NID:g17907664)"
4224..4246
/rpc_family="AT_rich"
5266..5556
/rpc_family="Alu"
5997..6061
/rpc_family="CR1"
6244..6318
/note="match to EST AL537636 (NID:g12801129)"
6244..6318
/note="match to EST AW814670 (NID:g17907664)"
6244..6318
/note="match to EST BB867724 (NID:g10316500)"
6547..6628
/rpc_family="L2"
6787..6814
repeat_region
```

```
repeat_region /rpc_family="(TTCA)n"
6802..6873
/rpc_family="L2"
repeat_region 6918..7214
/rpc_family="Alu"
repeat_region 7234..7340
/rpc_family="MIR"
repeat_region 7499..7802
/rpc_family="Alu"
7961..8399
/note="match to EST W25250 (NID:g1303104) zb68e09.r1"
misc_feature 8787..9076
/note="similar to Bos taurus EST BE653485 (NID:g11918617)"
misc_feature 8926..9076
/note="match to EST AL537636 (NID:g12801129)"
misc_feature 8926..9076
/note="match to EST AW814670 (NID:g17907664)"
misc_feature 8926..9076
/note="match to EST BB867724 (NID:g10316500)"
8999..9076
/note="similar to Sus scrofa EST BE232128 (NID:g9016846)"
9173..9233
/rpc_family="MaLR"
repeat_region 9234..9420
/rpc_family="MER1_type"
9421..9897
/rpc_family="MaLR"
9698..9699
/note="match to EST N95341 (NID:g1267676) zb68f09.s1"
repeat_region 9935..10207
/rpc_family="L1"
repeat_region 10879..11046
/rpc_family="MIR"
11228..11626
/note="match to EST BE155794 (NID:g11050977)"
11231..11905
/note="match to EST AL537636 (NID:g12801129)"
11231..11710
/note="similar to Sus scrofa EST BE232128 (NID:g9016846)"
11231..11566
/note="match to EST AW814670 (NID:g17907664)"
11231..11558
/note="match to EST BB867724 (NID:g10316500)"
11231..11445
/note="similar to Bos taurus EST BE653485 (NID:g11918617)"
11412..11967
/note="match to EST N95341 (NID:g1267676) zb68f09.s1"
11617..11963
/note="match to EST AL347484 (NID:g4084690) tc06e04.x1"
11742..12650
/note="match to EST AL577910 (NID:g12941490)"
11754..11994
/note="similar to Homo sapiens EST BE161875
(NID:g8624596)"
11902..12561
/note="similar to Homo sapiens EST BE739336
(NID:g10153317)"
misc_feature 2173..2188
/note="similar to Homo sapiens EST AL524174
(NID:g12787667)"
2352..2384
/rpc_family="L1"
2385..3107
/rpc_family="L1"
3117..3138
/rpc_family="AT_rich"
3120..3587
/rpc_family="MER1_type"
3588..3651
/rpc_family="MER1_type"
3877..3996
/note="match to EST BB867724 (NID:g10316500)"
3947..3996
/note="match to EST AL537636 (NID:g12801129)"
3956..3996
/note="match to EST AW814670 (NID:g17907664)"
4224..4246
/rpc_family="AT_rich"
5266..5556
/rpc_family="Alu"
5997..6061
/rpc_family="CR1"
6244..6318
/note="match to EST AL537636 (NID:g12801129)"
6244..6318
/note="match to EST AW814670 (NID:g17907664)"
6244..6318
/note="match to EST BB867724 (NID:g10316500)"
6547..6628
/rpc_family="L2"
6787..6814
repeat_region
```

Query Match 6.0%; Score 43.6; DB 9; Length 171265;
Best Local Similarity 56.0%; Pred. No. 0.34;
Matches 102; Conservative 0; Mismatches 79; Indels 1; Gaps 1;

```
QY 542 TTCGATGAGAGATTAGCAAGGGAAGAAAGAACTATGCGTAGCGGAGTTGTTTTCAT 601
DB 56108 TTTAGTAAAAAAGGAGAAATGCTGGGTTGTGTTACTCTGTTTTCAGCA 56049
QY 602 CATATAAAAAGGAGAAATGTTGCTCATATGTCAGATTTTGGGACGCTTTA-ACITT 660
DB 56048 GCTATTAAGGAGAAATTAATTCACACATCTTCACATCTTGGGAGGAAAAATTAAT 55989
QY 661 TATTGACAGGACTATCAATATATACAGATATTTGCAAAAAAAGAACTAATATATA 720
DB 55988 TCTGAGCTAGGCGCTCTCAAGGCTTTAGAGAGGAGAGAAACAAAGATGACAAAGAA 55929
```

Qy 721 CA 722
Db 55928 CA 55927

RESULT 11
AC012052/c
LOCUS
DEFINITION
AC012052 208729 bp DNA linear HTG 04-MAY-2001
Homo sapiens chromosome 2 clone RP11-337F19 map 2, WORKING DRAFT
SEQUENCE, 23 unordered pieces.

AC012052
AC012052.3 GI:8576227
HTG, HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 208729)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 2, clone RP11-337F19
Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 208729)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Balwin,J., Barna,N., Beckert,R., Boguslavsky,I., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domingo,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heathford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatsis,A., Klein,D.,
Lehoczky,J., Lieu,C., Locke,K., MacDonald,P., Margulis,N.,
McEwan,P., McGuirk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tittrell,A., Vassiliev,H., Vo,A., Wheeler,D., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT
Submitted (19-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 21, 2000 this sequence version replaced gi:7321517.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center Project name: L1185
Center clone name: 337_F_19

Summary Statistics
Sequencing vector: M13; M77815, 100% of reads
Chemistry: Dye-primer-amerham; 3% of reads
Chemistry: Dye-terminator Big Dye; 97% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 191830 bases at least Q40
Consensus quality: 199935 bases at least Q30
Consensus quality: 203385 bases at least Q20
Insert size: 210000; agarose-fp
Insert size: 206529; sum-of-ctrls
Quality coverage: 4.0 in Q20 bases; agarose-fp
Quality coverage: 4.1 in Q20 bas.

NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1080: contig of 1080 bp in length

1081 1180: gap of 100 bp
* 1181 2268: contig of 1088 bp in length
* 2269 2368: gap of 100 bp
* 2369 3746: contig of 1378 bp in length
* 3747 3846: gap of 100 bp
* 3847 6128: contig of 2282 bp in length
* 6129 6229 7907: contig of 1679 bp in length
* 7908 8007: gap of 100 bp
* 8008 10148: contig of 2141 bp in length
* 10149 10248: gap of 100 bp
* 10249 12511: contig of 2363 bp in length
* 12612 12711: gap of 100 bp
* 12712 14921: contig of 2210 bp in length
* 14922 15021: gap of 100 bp
* 15022 18148: contig of 3127 bp in length
* 18149 18248: gap of 100 bp
* 18249 22510: contig of 4262 bp in length
* 22511 22610: gap of 100 bp
* 22611 27371: contig of 4761 bp in length
* 27372 27471: gap of 100 bp
* 27472 32405: contig of 4934 bp in length
* 32406 32505: gap of 100 bp
* 32506 38167: contig of 5662 bp in length
* 38168 38267: gap of 100 bp
* 38268 46372: contig of 8105 bp in length
* 46373 46472: gap of 100 bp
* 46473 56551: contig of 10079 bp in length
* 56552 56651: gap of 100 bp
* 56652 64715: contig of 8064 bp in length
* 64716 64815: gap of 100 bp
* 64816 75459: contig of 10644 bp in length
* 75460 75559: gap of 100 bp
* 75560 90744: contig of 15185 bp in length
* 90745 90844: gap of 100 bp
* 90845 107238: contig of 16394 bp in length
* 107239 107338: gap of 100 bp
* 107339 128875: contig of 21537 bp in length
* 128876 128975: gap of 100 bp
* 128976 152422: contig of 23447 bp in length
* 152423 152522: gap of 100 bp
* 152523 178409: contig of 25887 bp in length
* 178410 178509: gap of 100 bp
* 178510 208729: contig of 30220 bp in length.

FEATURES
source
1..208729
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-337F19"
/clone_idb="RPC1-11 Human Male BAC"
1..1080
/note="assembly_fragment"
1181..2268
/note="assembly_fragment"
clone end:17
vector side:right"
2369..3746
/note="assembly_fragment"
3847..6128
/note="assembly_fragment"
6229..7907
/note="assembly_fragment"
8008..10148
/note="assembly_fragment"
10249..12611
/note="assembly_fragment"
12712..14921
/note="assembly_fragment"
15022..18148
/note="assembly_fragment"

```

misc_feature      18249..22510
                  /note="assembly_fragment"
misc_feature      22611..27371
                  /note="assembly_fragment"
misc_feature      27472..32405
                  /note="assembly_fragment"
misc_feature      32506..38167
                  /note="assembly_fragment"
misc_feature      38268..46372
                  /note="assembly_fragment"
misc_feature      46473..56551
                  /note="assembly_fragment"
misc_feature      56652..64715
                  /note="assembly_fragment"
misc_feature      64816..75459
                  /note="assembly_fragment"
misc_feature      75560..90744
                  /note="assembly_fragment"
misc_feature      90845..107238
                  /note="assembly_fragment"
misc_feature      107339..128875
                  /note="assembly_fragment"
misc_feature      128976..152422
                  /note="assembly_fragment"
                  clone_end:SP6
vector_side:left"
152523..178409
                  /note="assembly_fragment"
178510..208729
                  /note="assembly_fragment"

ORIGIN
Query Match      6.0%; Score 43.6; DB 2; Length 208729;
Best Local Similarity 56.0%; Pred. No. 0.35; Indels 1; Gaps 1;
Match 102; Conservative 0; Mismatches 79;

QY      542 TTGCATGAGAGATTAGCAAGCGAATAAATCATGCTAGCTGGAGTTGTTTCAAT 601
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      73634 TTTAGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 73575

QY      602 CATATAAAGGAGAAATTGTTGCTCATACTGTGACAGTTTCTGGAGCTTTA-AC 660
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      73574 GCTATPAAAAAGGAGAAATTATATCATCACACATCTCACACATTCGGAGGAAAAATAA 73515

QY      661 TATTCAGAGAGATATCAATCTATACAGATATGTCAAAAAAATAAAGACTAATAATA 720
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      73514 TCTGAGCTAGGCTCTCAAGGCTCTTAGAGACGAGAGAGAAACAAAGATGACAAAGAA 73455

QY      721 CA 722
        ||
Db      73454 CA 73453

RESULT 12
AC004553/c      80659 bp      DNA      linear      PRI 25-MAY-2002
LOCUS      Homo sapiens X GSHB-223P11 (Genome Systems Human BAC library)
DEFINITION      complete sequence.
ACCESSION      AC004553
VERSION      AC004553.1
KEYWORDS      GI:3451332
SOURCE      HTG.
ORGANISM      Homo sapiens (human)
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
                1 (bases 1 to 80659)
                Muzny,D., Aronson,A.D., Adams,C., Brundage,E., Bunce,C.,
                Carvelli,K., Chacko,J., Chen,J., Di,W., Ding,Y., Dugan,S.,
                Durbin,J., Forcum,J., Ganesh,R., Garcia,C., Goodman,M.,
                Gorrell,J.H., Haywood,M., Hernandez,J., Jackson,L., Jin,S.,
                Kampel,R., Karpathy,S., Kovar,C., Leal,B., Li,Y., Licharge,O.,
                Liu,W., Logan,O., Lu,J., Ly,T., Martinez,C., Oswal,G., Perez,L.,
                Rashid,N.D., Rowland,K., Savage,L., Scherer,S.E., Shen,H.,

```

```

TITLE      Simon,M., Stovall,K., Timms,K.M., Todd,J., Vo,Q., Williamson,A.,
            Worley,K.C., Yu,W., Chinnault,C., Nelson,D. and Gibbs,R.A.
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 80659)
AUTHORS      Worley,K.C.
TITLE      Direct Submision
JOURNAL      Submitted (11-APR-1998) Molecular and Human Genetics, Baylor
            College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
            3 (bases 1 to 80659)
REFERENCE      Worley,K.C.
            Direct Submision
            Submitted (25-AUG-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            4 (bases 1 to 80659)
REFERENCE      Worley,K.C.
            Direct Submision
            Submitted (17-SEP-1998) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            5 (bases 1 to 80659)
REFERENCE      Worley,K.C.
            Direct Submision
            Submitted (29-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            6 (bases 1 to 80659)
REFERENCE      Worley,K.C.
            Direct Submision
            Submitted (25-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            On Aug 25, 1998 this sequence version replaced gi:3402643.
            Sequencing is completed to a minimum standard of double strand
            coverage with a minimum of 2 clones and 2 reads with no ambiguities
            or 2 chemistries with a minimum of 2 clones and 3 reads with no
            ambiguities. If the sequence quality does not meet this standard,
            it will be indicated in the annotation.

COMMENT      The repeat regions shown were identified using RepeatMasker by
            Adrian Smit.
            Sequence similarities were identified using Powerblast by Jinghui
            Zhang.
            Exon/Intron boundaries of identified genes were chosen if there
            were canonical splice junctions that maintained sequence continuity
            across the splice junctions.
FEATURES      source
            1..80659
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="X"
                /clone="GSHB-223P11"
            1..1999
                /note="overlaps bases 114570..116568 of clone AC112492"
                /function="clone overlap"
            1310..1350
                /rpt_family="L2"
            1527..1639
                /rpt_family="MIR"
            complement(1641..1930)
                /rpt_family="AluY"
            1933..2050
                /rpt_family="MIR"
            complement(3587..3670)
                /rpt_family="MIR"
            3679..3861
                /rpt_family="MER53"
            4687..4829
                /rpt_family="MER5B"

```

```
repeat_region complement(5319,.5358)
/rpt_family="AT_rich"
repeat_region complement(5853,.5898)
/rpt_family="AT_rich"
repeat_region complement(5943,.6244)
/rpt_family="MLT1B"
repeat_region complement(6228,.6290)
/rpt_family="MLT1D"
repeat_region complement(6354,.6624)
/rpt_family="L1MA2"
repeat_region complement(6631,.6758)
/rpt_family="MLT1D"
repeat_region complement(7635,.7690)
/rpt_family="(CA)n"
repeat_region 7833,.8023
/rpt_family="MER3"
repeat_region 8188,.8352
/rpt_family="MERSA"
repeat_region 9035,.9057
/rpt_family="AT_rich"
repeat_region 9148,.9227
/rpt_family="MIR"
repeat_region complement(9246,.9302)
/rpt_family="AT_rich"
repeat_region complement(9663,.10244)
/rpt_family="L1MC4"
repeat_region complement(10630,.10701)
/rpt_family="(TA)n"
repeat_region complement(11161,.11220)
/rpt_family="(GA)n"
repeat_region complement(12725,.12846)
/rpt_family="(TA)n"
repeat_region complement(12853,.12932)
/rpt_family="(CATA)n"
repeat_region 13300,.13428
/rpt_family="MIR"
repeat_region complement(13783,.14150)
/rpt_family="L1PA8"
repeat_region 14162,.14188
/rpt_family="AT_rich"
repeat_region complement(14686,.14806)
/rpt_family="FLAM_C"
repeat_region 15736,.15925
/rpt_family="L2"
repeat_region complement(16745,.17110)
/rpt_family="L1M4"
repeat_region 17480,.17941
/rpt_family="MLT1C"
repeat_region 18083,.18134
/rpt_family="MADE1"
repeat_region complement(18177,.18784)
/rpt_family="L1MA9"
repeat_region 19560,.19858
/rpt_family="AluSg"
repeat_region complement(20203,.20229)
/rpt_family="AT_rich"
repeat_region 20350,.20551
/rpt_family="MER4_internal"
repeat_region 20563,.21094
/rpt_family="MER4_internal"
repeat_region 21096,.21246
/rpt_family="MER4A2"
repeat_region 21247,.21510
/rpt_family="MER4A"
repeat_region 21804,.22111
/rpt_family="AluSp"
repeat_region complement(22905,.22958)
/rpt_family="MIR"
repeat_region 23531,.24248
/rpt_family="L1ME3A"
repeat_region complement(24998,.25210)
/rpt_family="L1PA8"
repeat_region 25211,.25359
```

```
repeat_region /rpt_family="MER11A"
25357,.26294
/rpt_family="MER11B"
repeat_region complement(26295,.26395)
/rpt_family="L1PA8"
repeat_region 26608,.26927
/rpt_family="AluSg"
repeat_region 27578,.27878
/rpt_family="AluSx"
repeat_region complement(28196,.28279)
/rpt_family="L2"
repeat_region complement(28709,.28748)
/rpt_family="(TAA)n"
repeat_region complement(28749,.29386)
/rpt_family="L1PA4"
STS 29091,.29222
/standard_name="G19948"
/db_xref="dbSTS:32826"
repeat_region 29361,.29554
/rpt_family="L1P"
repeat_region complement(30273,.31278)
/rpt_family="MER11A"
repeat_region complement(32126,.32237)
/rpt_family="(GAA)n"
STS 32378,.32676
/standard_name="DXS67"
complement(33176,.33223)
/rpt_family="MIR"
repeat_region 33998,.34051
/rpt_family="MIR"
repeat_region complement(35042,.35071)
/rpt_family="AT_rich"
repeat_region complement(35418,.35439)
/rpt_family="AT_rich"
repeat_region complement(38090,.38123)
/rpt_family="AT_rich"
```

Query Match 5.9%; Score 42.8; DB 9; length 80659;
Best Local Similarity 51.0%; Pred. No. 0.52;
Matches 101; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```
QY 517 GAACATAGTGAATAGGATGTAACCTTTCGATGAGAGATTCGACAGCGGAAAAAATCT 576
|||||
DB 33326 GAACATAGTGTCTCAAGAGTTCGAACTCCGAGTTCGAACTGATGTTTGAATAATPA 33267
|||||
QY 577 GGCTTAGCTGGAGTGTGTTTCAATCATATATAAGGAGAAATGTCGACTATGTA 636
|||||
DB 33266 AACATGCTAAGAGATTAATTAAGTGTATTTTAAGGAAATTAATTAAGGCTTCTGTC 33207
|||||
QY 637 CAGTTTCTGGACGCTTAACCTTTATTCAGAGGACTATCAATCATACAGATATTGTC 696
|||||
DB 33206 CAGTTTCTCATCAACCAAAAGAGTATATAAAAAAATTAATTCGACATATTGTG 33147
|||||
QY 697 AAAAAAAAAAAGACTAA 714
|||||
DB 33146 ATGAAAAAAAAAAAAA 33129
|||||
```

RESULT 13

AC021710 195932 bp DNA linear HTG 04-APR-2000
LOCUS Homo sapiens chromosome X clone RP11-12D5 map X, WORKING DRAFT
DEFINITION SEQUENCE, 9 unordered pieces.
ACCESSION AC021710
VERSION AC021710.4 GI:7408016
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome X, clone Rpl1-12D5

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 195932)
Birten, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckert, R., Beda, F.,
Boguslavsky, L., Bouknight, B., Brown, A., Burkett, G., Castle, A.,
Choe, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearlino, K., Dewar, K., Domino, M., Doyle, M., Fennestor, J.,
Ferrera, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J. C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,
Lander, E., Lehotzky, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGuck, A., McKernan, K.,
McPherson, R., Meldrum, J., Menus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirelli, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A., and Zody, M.

TITLE
JOURNAL

Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 4, 2000 this sequence version replaced gi:16939568.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L5924

Center clone name: 12_D_5

Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731

Consensus quality: 189137 bases at least Q40

Consensus quality: 192288 bases at least Q30

Consensus quality: 193686 bases at least Q20

Insert size: 190000; agarose-fp

Insert size: 195132; sum-of-contigs

Quality coverage: 5.0 in Q20 bases; agarose-fp

Quality coverage: 4.9 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
consists of 9 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1279: contig of 1279 bp in length
* 1280 1379: gap of 100 bp
* 1380 10400: contig of 9021 bp in length
* 10401 10500: gap of 100 bp
* 10501 19494: contig of 8994 bp in length
* 19495 19594: gap of 100 bp
* 19595 32680: contig of 13086 bp in length
* 32681 32780: gap of 100 bp
* 32781 49150: contig of 16370 bp in length
* 49151 49250: gap of 100 bp
* 49251 73382: contig of 24132 bp in length
* 73383 73482: gap of 100 bp
* 73483 104624: contig of 31142 bp in length
* 104625 104724: gap of 100 bp
* 104725 142393: contig of 37669 bp in length
* 142394 142494: gap of 100 bp
* 142494 195932: contig of 53439 bp in length.
Location/Qualifiers
1. 195932

FEATURES
source

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="RP11-1205"
/clone_1b="RP11-11 Human Male BAC"
1. 1279
/note="assembly_fragment"
1380. 10400
/note="assembly_fragment"
10501. 19494
/note="assembly_fragment"
19595. 32680
/note="assembly_fragment"
32781. 49150
/note="assembly_fragment"
clone_end:17
vector_side:right"
49251. 73382
/note="assembly_fragment"
73483. 104624
/note="assembly_fragment"
104725. 142393
/note="assembly_fragment"
142494. 195932
/note="assembly_fragment"
clone_end:8P6
vector_side:left"

ORIGIN

Query Match 5.9%; Score 42.8; DB 2; Length 195932;
Best Local Similarity 51.0%; Pred. No. 0.61; Mismatches 97; Indels 0; Gaps 0;
Matches 101; Conservative 0;

QY 517 GAACATAGTATAGGATGTAACCTTCGATGAGAAATTCACGAGAAAACTAT 576
DB 188972 GAACATAGTATGCAAGATCAAACTCGAGTTCAAGTTAGCATGTTTGGAAATTA 189031
QY 577 GCGTACTGCGAGCTGTTTTCATCATATATAAAGGAGAAATTTGCTCACTATGTGA 636
DB 189032 AACATGCTTAAGATTAACCTTAAGTGTATTTTAAAGGAATTAACCTTGAGCGTTCTGTGC 189091
QY 637 CAGTTTCGCGAGCGTTTAACTTTATTCGACGAGCTCAATCATCAGATATTGTC 696
DB 189092 CAGTTTCCTCATACCAAAAGAGATATATAAAAAACAATTAATGCACATCATGTG 189151
QY 697 AAAAAAAAAAAGACTTA 714
DB 189152 ATGAAAAAAAAAAAAA 189169

RESULT 14
AL359542/c 28858 bp DNA linear PRI 09-MAR-2001
LOCUS Human DNA sequence from clone RP6-190D15 on chromosome Xq25-26.1,
DEFINITION complete sequence.
ACCESSION AL359542
VERSION AL359542.13 GI:13274757
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 28858)

AUTHORS Heach, P.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2001) Sanger Centre, Hinxton, Cambridgehire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk

COMMENT On Mar 12, 2001 this sequence version replaced gi:13092292.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emu, ENML, Swi, SWISSPROT, Tr, TREMBL, Wp, WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormp This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChX> RP6-190D15 is from the library RPCT-6 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pPAC4
IMPORTANT: This sequence is not the entire insert of clone RP6-190D15. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP4-537K23 is at 28759 in this sequence.
The true right end of clone RP4-753P9 is at 100 in this sequence.

FEATURES

```
source
1..28858
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="q25-26.1"
/clone="RP6-190D15"
/clone_1lb="RPCT-6"
39..338
/notes="AluY repeat: matches 1..300 of consensus"
347..646
/notes="AluX repeat: matches 1..299 of consensus"
3076..3129
/notes="27 copies 2 mer tc 77% conserved"
5914..6224
/notes="AluYB8 repeat: matches 1..311 of consensus"
6287..6596
/notes="AluSq repeat: matches 1..312 of consensus"
6597..8469
/notes="L1MB8 repeat: matches 4267..6133 of consensus"
8470..8757
/notes="AluX repeat: matches 5..302 of consensus"
8758..8796
/notes="L1MB8 repeat: matches 6133..6171 of consensus"
8818..9104
/notes="AluSg1 repeat: matches 1..289 of consensus"
9345..9465
/notes="L12 repeat: matches 2602..2728 of consensus"
10287..10362
/notes="38 copies 2 mer aa 69% conserved"
10380..10462
/notes="L1MB3 repeat: matches 6035..6127 of consensus"
10464..10518
/notes="7SK repeat: matches 1..53 of consensus"
10835..11133
/notes="AluX repeat: matches 1..298 of consensus"
11889..12080
/notes="AluSg repeat: matches 24..296 of consensus"
12081..12122
/notes="AluSg repeat: matches 253..294 of consensus"
12927..12974
/notes="24 copies 2 mer gt 72% conserved"
13691..13842
/notes="MIR repeat: matches 65..216 of consensus"
13957..14317
/notes="L1M4 repeat: matches 4795..5162 of consensus"
14354..14450
/notes="L1MA2 repeat: matches 6213..6308 of consensus"
14610..14726
/notes="L1MA9 repeat: matches 6110..6218 of consensus"
14732..14837
/notes="MERB repeat: matches 49..156 of consensus"
14838..14908
/notes="TIGGER2 repeat: matches 1..70 of consensus"
14909..15125
/notes="MERB repeat: matches 1..345 of consensus"
15171..15249
/notes="TIGGER2 repeat: matches 1..73 of consensus"
15250..15315
/notes="33 copies 2 mer aa 68% conserved"
15620..15923
/notes="AluSq repeat: matches 1..312 of consensus"
15943..15956
/notes="L1MB4 repeat: matches 5639..5652 of consensus"
15957..16210
/notes="AluSg repeat: matches 38..290 of consensus"
16081..16168
/notes="Weakly double-stranded"
16213..16724
/notes="L1MB4 repeat: matches 5646..6176 of consensus"
16725..17110
/notes="MLT1B repeat: matches 1..388 of consensus"
17111..17277
/notes="L1MB4 repeat: matches 5487..5639 of consensus"
17640..17949
/notes="AluY repeat: matches 1..309 of consensus"
18445..18745
/notes="AluX repeat: matches 1..302 of consensus"
19047..19348
/notes="AluB repeat: matches 1..291 of consensus"
19523..19744
/notes="L1PA12 repeat: matches 5937..6158 of consensus"
19745..20039
/notes="L1P4 repeat: matches 5345..5653 of consensus"
20040..20339
/notes="AluY repeat: matches 1..301 of consensus"
20391..20689
/notes="AluSg repeat: matches 1..299 of consensus"
20689..21369
/notes="L1PA8 repeat: matches 5489..6163 of consensus"
21430..21459
/notes="U6 repeat: matches 1..30 of consensus"
21892..22025
/notes="MIR repeat: matches 115..254 of consensus"
23037..23338
/notes="AluX repeat: matches 1..296 of consensus"
23378..23562
/notes="L1M4 repeat: matches 5627..5818 of consensus"
23770..23985
/notes="AluX repeat: matches 1..216 of consensus"
23986..24027
/notes="Alu repeat: matches 259..298 of consensus"
24570..24787
/notes="L1MB repeat: matches 5500..5714 of consensus"
25661..25968
/notes="AluX repeat: matches 1..312 of consensus"
25970..26067
/notes="49 copies 2 mer ag 61% conserved"
26552..26865
/notes="AluX repeat: matches 1..312 of consensus"
27019..27089
/notes="MER51B repeat: matches 453..526 of consensus"
27090..27396
/notes="AluSg repeat: matches 1..306 of consensus"
27397..27415
/notes="MER51B repeat: matches 525..543 of consensus"
```



```
repeat_region 27925..28003
                /note="L2 repeat: matches 2535. .2736 of consensus"
repeat_region 28662..28805
                /note="MIR repeat: matches 110. .256 of consensus"
ORIGIN
```

This Page Blank (uspto)